SOLAR RADIO BURST EFFECTS ON GLOBAL POSITIONING SYSTEM RECEIVERS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alessandro Paolo Cerruti

January 2008

© 2008 Alessandro Paolo Cerruti

ALL RIGHTS RESERVED

SOLAR RADIO BURST EFFECTS ON GLOBAL POSITIONING SYSTEM RECEIVERS

Alessandro Paolo Cerruti, Ph.D.

Cornell University 2008

This thesis presents a series of studies investigating solar radio burst effects on Global Positioning System (GPS) receivers along with supporting instrumentation and analysis techniques. Solar radio bursts are a space weather phenomenon with its origins at the sun. Although solar radio bursts have been widely studied since the 1960s and have been known to cause notable problems for radio communication systems such as cell phone networks (*Lanzerotti et al.*, 1999) their impact on GPS was underestimated and largely ignored by the community. Recent large solar radio burst events in conjunction with accurate carrier-to-noise measurements from GPS receivers has allowed for the first precise qualitative and quantitative analysis of their impact. To the receiver, a solar radio burst is a wide-band radio interference source that causes an effective decrease in the received carrier-to-noise ratio. The analysis of moderate events on 7 September 2005 allowed for the prediction that larger solar radio bursts would present a significant challenge to GPS availability as soon as the next solar maximum in 2011-2012 (*Cerruti et al.*, 2006). The future came sooner than expected when the record setting solar radio burst of 6 December 2006 caused wide-spread outages of GPS receivers. The event exceeded 1,000,000 SFU, was about ten times larger than any previously reported event, and was all the more surprising since the solar radio bursts occurred near solar minimum. These events had a drastic impact on several critical GPS systems utilized by the scientific community, the Federal Aviation Administration, oil-rig operations,

orbiting satellites, and surveying. The size of the December 2006 bursts strongly suggests that the historical record may be inaccurate and raises the possibility for even more intense solar radio bursts during the next solar maximum that will significantly impact the operation of GPS receivers.

BIOGRAPHICAL SKETCH

Alessandro (Alex) Paolo Cerruti was born 22 March 1980 in Wimbledon, United Kingdom to Alberto and Paolo Cerruti, both of Biella, Italy. He emigrated to the United States of America in November 1982 and grew up in Fairfield, Connecticut where he played with his Legos, dreamed of building and flying airplanes, and attended the public school system. In 1992 he was proud to become an American Citizen, and in the fall of 1998 he attended Cornell University to pursue an interest in Electrical Engineering. He completed his Bachelor of Science in Electrical and Computer Engineering one semester early, and due to the technology-bubble bust and the tragic events of September 11, 2001, he decided to stay one more year for his Master of Electrical and Computer Engineering. While fulfilling his obligations for his Master of Engineering, Alex was hired as a Teaching Assistant for his favorite undergraduate course: ECE 415, Global Positioning System Theory and Design. To further his knowledge in Global Positioning Systems (GPS), he joined the GPS group and began working with Professors Paul Kintner and Mark Psiaki. GPS was, in a way, fulfilling his childhood aviation dreams. He decided to pursue a career in the field, so in January 2003 Alex began his Ph.D. with Professor Paul Kintner. By a fortunate series of events, and because he was tasked to write a research paper for Professor Donald Farley's Incoherent Scatter Radar course, Alex stumbled upon the first quantitative measurements of solar radio burst effects on GPS. Through Professor Kintner's encouragement, he further explored the ramifications of solar radio bursts on GPS and the historic bursts of December 2006. Alex is looking forward to finally moving out of Ithaca (after all he has been here for over nine years now) and further pursuing his interests in GPS.

Photograph of the author as requested by the Special Committee.

Dedicated to my family and friends whose faith and support has sustained me through this odyssey.

ACKNOWLEDGEMENTS

I have many individuals and organizations to whom an acknowledgement is owed. Without their help and support, this Ph.D. would have been an insurmountable task.

I thank my advisor, Dr. Paul Kintner, for having patience with me, for believing that I had the right stuff, and for being a friend. Paul has provided wonderful guidance without being managerial or overwhelming, thus allowing me to explore my interests at my pace. Paul encouraged me to determine the cause of the bizarre signal fades I was observing in the GPS data from Arecibo that led to the first quantitative measurements of solar radio burst induced carrier-to-noise ratio fades on GPS. Paul was responsible for organizing the original press release through the Cornell University Press, as well as the press release through NOAA in April 2007. Both were experiences that I shall never forget. I am grateful to him for giving me the opportunity to teach his course (ECE 415, Global Positioning System Theory and Design) while he was off at various conferences and meetings, and for creating ECE 584, Advanced GPS Receiver Design, from the GPS software receiver that myself and other Master of Engineering students had developed.

Dale Gary of the New Jersey Institute of Technology provided valuable solar radio burst data from the OVSA. Conversations in person and over email were always enjoyable, insightful, and led to many 'Ah ha!' moments. I extend my deepest thanks to Anthea Coster of MIT Haystack for her excitement at the 2006 CEDAR conference in Santa Fe, NM when I originally showed her the solar radio burst data and the effects on GPS. Her wonderful reaction motivated me further, and she continues to be a wonderful colleague. Tony Mannucci and Robert Meier provided analysis of the impact on NASA's GDGPS network. They supplied insight into the broader effects of the solar radio burst on GPS which was a valuable

contribution to the NOAA press release and to this thesis. I also thank Pat Doherty for her help with the WAAS data and analysis, which most of this thesis relies upon. You have all provided a key contribution to my success and I thank you.

I extend my gratitude to past and current member of the Space Physics and Engineering Group and Cornell University. Brent, you made my life hell at times, but it was all worth it in the end, thank you. I have learned so much from you, and you have been a wonderful friend and colleague. Jon Makela, Erik Klatt, and Ted Beach, thank you for your guidance and conversations on space physics. To all other students in the group, thank you, and good luck. I would also like to thank Laurie Shelton for proofing (affects of effects?) and for being my travel agent. Steven Powell is a remarkable engineer who can seemingly solve any problem. I have learned more practical knowledge from you than anyone else–after all, who else has the resourcefulness to launch turkey fryers and toilet floats into space? Finally, I thank Dr. Donald Farley, Dr. Wesley Swartz, and Dr. Charles Seyler for being wonderful and supportive instructors. Last but not least, I thank my Special Committee members Dr. David Hysell and Dr. Michael Kelley. Their guidance and suggestions were invaluable throughout the entire process.

Gratitude is extended to Mark Psiaki, who in many ways was a fourth Special Committee member. Mark's knowledge and insight has saved me from wasted time and grief. The parties you and Mary hosted were fantastic and will always be remembered. Thank you Todd Humphreys for countless conversations and mini-tutorials, I've learned so much from you.

I thank the Faculty at the School of Electrical and Computer Engineering at Cornell University, most of whom have had me as a student at some point over these nine years. I would especially like to thank John Belina for being a wonderful undergraduate advisor, for endless conversations, for giving me the opportunity to

teach, and for always encouraging me to explore. I extend my deepest gratitude to Dave Delchamps for his counsel and help. There are numerous individuals at the School of Electrical and Computer Engineering that I would like to thank including Scott Coldren, Dorothy Palladino, Cheryl Francis, Paul Schuh, Robert Beaver, Jason Seymour, Bill Mutch, Patty Gonyea, Sandi Goodwin, and Jamie Dal Cero. All have made my time here easier in some respect. I am in debt to Theoretical and Applied Mechanics and Wolfgang Sachse for supporting me financially during my first semester as a Ph.D. student at Cornell University.

There are many people and institutions who have helped me while I was a Ph.D. student or recognized my achievements. Eurico De Paula coordinated three research trips to Brazil and has been a wonderful colleague. I am indebted to the Instituto Nacional de Pesquisas Espaciais and their employees for hosting our research trips to the country. Their hospitality helped the trips run flawlessly. I also appreciate the Brazilian Federal Police for not arresting Brent and I in São Luis in 2003. I thank NOAA and in particular Joe Kunches, Carmeyia Gillis, Bill Murtagh, and Brigadier General David L. Johnson, USAF (Ret.) for their support and help with the April 2007 press conference. I acknowledge Fugro Chance and Richard Barker for their analysis on the solar radio burst events of December 2006. A thank you is extended to the NAIC, the Arecibo Observatory, and in particular Hien Vo and Sixto Gonzales for their help in setting up the GPS equipment that made the first observations of solar radio bursts. Finally, I would like to thank the CEDAR community as well as The Institute of Navigation for the awards and honors they have given me.

I extend my deepest gratitude and respect to Cathryn Mitchell and Gary Bust. I will never forget the time we have spent together and you have been more than just colleagues to me. I look forward to the next time we can play our Sh'em or

Sh'em? game.

My family has tolerated and supported me throughout the thesis process. Dad, your constant nagging was too much at times, but admittedly, kept me on task. Mom, your never-ending support helped me in ways that I cannot even begin to express. Sara and Giò, thank you for the visits to Houston that I always enjoyed, especially when Gaia and Sofia came along. Tony and Gail, thanks for calling to check up on me and sending ridiculous movies of Mia, which were always a welcome respite from the work. Nonna Pia, sono contento che in finendo questa tesi ti ho portato tanta felicità. I love you all.

Lastly, I'd like to thank my family at Cornell–that being my friends. Without you I would have never survived nine and a half years here. Nini, I will miss our ridiculous nights out dancing, our dinner dates, and our inane conversations. Te quiero mucho. Dan H., I know you will hate me for this, but I thank you from the bottom of my heart: you are wonderful. Dave, thanks for always being there and for our cooking nights. Ben, for endless IM conversations over half the globe. Matt, thanks for your advice and for taking care of Colter on numerous occasions. Erik L., your political incorrectness was always a great stress relief. Brady, thank you for taking the torch and running with it. Asti, I'm going to miss inadvertently scaring you and our crazy conversations. Mike N., thanks for all your help with homework and insight. Mark, you kept me sane, thank you. Mike S., I could always count on you. Bill, thanks for always supporting me. Andrew, I look forward to chilling with you again when I move to Boston. Thanks Jeremy and the Jasons for countless game and trivia nights. Jamie and Kim I owe you big time for the conversations and the hugs.

My graduate work at Cornell University was supported by the Office of Naval Research under grant number N00014-04-10105. I am grateful to the School of

Electrical and Computer Engineering at Cornell University for the Teaching Assistantships.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

The History of Navigation

One man's noise is another man's signal. - Unknown

1.1 Introduction

Proper navigation requires accurate knowledge of the size and shape of the earth and the ability to map it properly, a science known as geodesy. To accurately navigate from one place to the next, one must also have good knowledge of time, a science known as horology. Finally, astronomy allows the navigator to accurately determine where he is. Humans have navigated the globe for thousands of years and the tools of navigation took centuries to perfect. Accurate navigation only came about in the early 1800s. The following account on the history of navigation is based on *Sobel* (1995).

On earth, the globe is split into imaginary lines. The parallel lines that girdle the earth in an east-west direction are known as latitudes, whereas the lines that go north-south through the geographic pole are known as the longitudes. Since at least three centuries before the birth of Christ, the world has been divided into these lines of latitude and longitude.

Latitude has always been an easy quantity to measure because nature has provided natural guide-posts. The ancients studied the heavens and noted that the sun and moon pass nearly overhead at the equator. Furthermore, the northernmost and southernmost excursions of the sun marked the Tropic of Cancer and the Tropic of Capricorn. Finally, the Pole Star in the northern hemisphere is a natural

guidepost whereby latitude is easily determined.

Longitude, on the other hand, literally took at least two thousand years to master. Longitude is naturally difficult to ascertain as there exists no natural guide post. The passage of time enables one to determine his longitude. By keeping the known time of a reference longitude (e.g. the home port) and by updating a second clock to local noon gauged by the passage of the sun, a user can easily determine his longitude. Developing a clock accurate enough to maintain the home port time and hence measure longitude is no small feat. The timeline of navigation contains many important milestones and innovations that led to the mastery of longitude. Each innovation, however, had shortcomings that spurred additional technology and further improved the science. Determination of longitude is an important milestone that forever changed how humans travel the world.

1.2 The Quest for the Longitude

The first practical technique for determining longitude accurately was invented by Galileo Galilei. In 1610 he proposed a method of studying the predictable eclipses of the moons of Jupiter to determine longitude. The technique was widely accepted by cartographers on land and the technique led to the first accurate measurements of the speed of light by Roemer in 1676. The method, however, was only practical at night when Jupiter was visible. A rocking ship made it difficult to keep Jupiter within the field of view of a telescope, and so longitude at sea could not be determined.

The history of the seas abounds with stories of lost souls suffering of starvation, thirst, scurvy, and shipwreck. In one incident, over 2000 English sailors lost their

Figure 1.1: An E. Dent and Co. Liquid Boat Compass with Patented Binnacle (ca. 1890).

lives on 22 October 1707 when their commander Sir Clowdisley lost his bearings in a foggy sea and scattered his warships against the rocks of the Scilly Isles. Captains of the fifteenth, sixteenth, and seventeenth centuries were prone to such disasters because they relied on dead reckoning to measure longitude. Logs were thrown overboard to determine the ship's velocity. Measurements of direction were made from the heavens or from magnetic compasses, the passage of time on an hourglass or pocket watch. Factoring in the measurements, ocean currents, variable winds, and errors in judgement, the captains would determine their longitude. Too often the crude measurements would lead to dead seamen.

Captains had the option of taking well-known sea routes by following particular lines of latitude. The congestion along these waters guaranteed that pirates and

war-faring nations often encountered whaling and merchant ships. In 1592 near the Azores, six English men-of-war ambushed and pillaged the Portuguese galleon *Madre de Deus* and sailed away with a bounty worth a half-million pounds sterling– at that time, an amount worth half of the entire English Exchequer. By the late 1700s, a few hundred ships a year would sail from England to the West Indies, and the loss of even one of these ships resulted in significant monetary losses.

Mariners were in search of new routes to safely traverse the seas, but new routes required accurate determination of longitude. The shipwrecking and plundering of ships brought the need for the determination of longitude to the attention of the leaders of many nations, including King George III of England and King Louis XIV of France. Many prizes were offered for an accurate determination of the longitude, but the greatest prize offered was the Longitude Act of 1714.

In 1714, merchants and sailors petitioned for the formation of a committee to consider solving the longitude problem in a reasonable way. The seafarers requested a fund to support research and development and demanded that the inventor be rewarded grandly. In May 1714, a Parliamentary committee at Westminster Palace responded to the challenge and sought the advice of Sir Isaac Newton and Edmond Halley. Together they suggested that the only feasible means of determining longitude was either by keeping exact time or by astronomical observations. Finally, on 8 July 1714, Queen Anne issued the Longitude Act which provided the following prizes:

- £20,000 for a method to determine longitude to an accuracy of half a degree of a great circle
- £15,000 for a method accurate to within two-thirds of a degree
- £10,000 for a method accurate to within one degree

In addition to prizes, the Board of Longitude (comprised of scientists, naval officers, and government officials) could disburse funds to aid impoverished inventors who could bring promising ideas to fruition. To win a prize, the technique had to be tested on one of Her Majesty's ships on a voyage from Great Britain to any port in the West Indies while maintaining the appropriate accuracy for the prize.

Many impractical ideas were presented to the Board including improvements to ship's rudders, water purification techniques, modifications to ship's sails for stormy weather, perpetual motion machines, and proposals to square the circle or make sense of the value of pi. Accurate clocks were also proposed, but in 1714, clocks drifted by about 15 minutes per day. To win the top prize, a clock had to drift no more than two minutes (half a degree of longitude) over a six week period, the typical length of a trans-Atlantic voyage.

Clocks before the Longitude Act proved troublesome for several reasons. Galileo and Huygens both studied the pendulum and developed clocks based on its premise. Huygens understood the sensitivity of the pendulum to heat, which caused the length of the pendulum to change, to gravity, which caused the pendulum to speed up or slow down, and to the effect of friction and heat-sensitive lubricants. Furthermore, the pendulum is useless on rough seas where the rocking of the ship disrupts its natural motion. Huygens solution was a spiral balance spring as an alternative to the pendulum. Further advances of his invention were thwarted when Robert Hooke independently invented the spiral balance spring and concurrently lay claim to a patent for his invention from the English Royal Society. The patent feud never was resolved and neither Huygens nor Hooke produced a marine chronometer. Both regarded as highly successful in their fields, their failure to develop a marine chronometer seemed to dull the spirits of clockmakers and the

Board itself. Newton, an astronomer, and the Board became doggedly convinced that the solution lay in the stars, and not in mechanical clocks.

John Harrison (English, b. 24 March 1693, d. 24 March 1776) was a self-educated clockmaker. Much of his early life is not well known except that he acquired an immense knowledge by reading books and learned woodworking from his father. Around 1712, a clergyman allowed John to borrow a manuscript copy of a lecture series on natural philosophy by mathematician Nicholas Saunderson at Cambridge University. Armed with Saunderson's notes, Newtown's Principia, determination, and no knowledge of watchmaking, in 1713 John built his first pendulum clock made almost entirely of wood.

John showed remarkable insight into building stable clocks. For example, he avoided oils and metals that corroded for fear that they would degrade the time-piece's accuracy. He carefully selected woods for their grain patterns and growth rates to ensure their strength. For parts requiring smooth function, he selected lignum vitae, which exudes its own lubricating oils. As he moved on to metals, he began to fabricate pendulums of alternating strips of different metals to counteract the effects of temperature. John also developed a friction-free clock escapement mechanism. By 1727, John had developed pendulum clocks that could keep time accurate to one second over the course of a month. Armed with friction-free escapements and knowledge to counteract temperature effects, John turned his efforts towards making a clock that was sea-worthy and winning the Board's prize.

John recognized, however, that pendulum clocks could not keep time on a rocking ship, so he developed a counter-balanced see-sawing mechanism that could keep time regardless of its tilt. By 1730, John traveled to London to visit the Board.

Tired of mediocre attempts, the Board would not entertain John's ideas, so instead he visited Halley at the Royal Observatory. Halley was impressed by John's clock but knew that the Board would not entertain a mechanical solution when they were stubbornly expecting a solution from the astronomers. Halley sent John to visit watchmaker George Graham, who was duly impressed with John's drawings and dispatched him with a loan to build his first marine chronometer, the H-1.

Built in five years, the H-1 was not initially presented to the Board, but rather to the Royal Society. With the Society's endorsement, the H-1 endured a week-long sea trial to Lisbon aboard the *Centurion*, and returned on a month long voyage through gales and calms aboard the H.M.S. *Orford*. The accuracy of John's clock so impressed the commander of the *Orford*, that he sent John back to London with a certificate of praise. The Board, hearing of John's clock, convened for the first time since its creation to marvel over his seemingly-unreal contraption. Although the clock never lost more than a few seconds over the entire voyage, John was critical of his time-piece. Rather than attempt to claim the prize, he insisted that the Board loan him money for improvements. So they did, and so he returned with the H-2 in 1737.

The H-2 (a huge device weighing in at 86 pounds) boasted several innovations which further improved its accuracy including more responsive temperature compensation devices and a mechanism to ensure a uniform drive. The machine passed all tests that the Royal Society could muster, including heating, cooling, and agitation far worse than any ship at sea could encounter. The Society gave its full backing of the machine including an assertion that "...the motion is sufficiently regular and exact, for finding the Longitude of a Ship within the nearest Limits proposed by Parliament and probably much nearer." Harrison, however,

was not happy. He disappeared for twenty years this time to build the H-3, and in the meantime fell prey to the Board's adamant belief that the solution to finding longitude lay in the stars.

While John Harrison was tinkering with and improving his marine chronometers, several innovations in astronomical tools came about that forever changed the art of navigation and proved to be a serious competitor to John's clocks. In 1731, about the time when John received the loan from Graham for his H-1, two inventors independently created an instrument which could accurately determine the distance between two celestial bodies as well as their elevation, regardless of the rocking of the ship. John Hadley's Quadrant (an American named Thomas Godfrey was the other inventor) was the device that finally allowed the clockworks of heaven to be used in determining longitude. Soon thereafter the sextant was invented which allowed for more accurate measurements. The sextant and similar devices enabled the method of Lunars to determine longitude.

Taking Lunars proved to be quite difficult. A navigator had to measure the angular distances and elevations of the moon and stars or the sun to a high precision. Atmospheric refraction would complicate measurements near the horizon and had to be accounted for. Furthermore, the tables from which the user's position was ascertained were tabulated from the center of the earth, so the navigator had to take into account lunar parallax. The end result was that the navigator had to be a rather good mathematician and astronomer. Although the tables were cheap to obtain, the calculations and experience required by the navigator were significant.

By the time Lunars was a proven technique for determining longitude, Harrison had completed the H-3, deemed that it could be resized to a more manageable device, and created the H-4. The H-4, presented to the Board in 1760, was only

5 inches in diameter and weighed a scant 3 pounds. In comparison to Lunars, the H-4 provided an easy solution for determining longitude, but the clock was quite expensive. His skeptical opponents did not believe it was possible for a timepiece to maintain the accuracy required, and as a result he was subjected to many trials beyond the original requirements set forth by the Board.

His main competitor was the Reverend Nevil Maskelyne. Maskelyne was a fervent advocate for the Lunars method. He was obstinately convinced the method of Lunars was superior and John's clocks were just a lucky fluke. He did everything within his power (even unfairly) to see that John's clock would not win the Longitude Act.

In the end, John Harrison nor any other individual actually won the Longitude Act. Intervention by King George III finally saw Harrison recognized with a total disbursement of £18,750, but the prize was disbursed by parliament and was not the coveted prize of the Board. The Board of Longitude disbursed over £100,000 to various individuals in their 114 year history, but never once awarded the grand prize.

Regardless, John's clocks proved their worth, and were hailed by captains the world over, including lavish praises from James Cook. Although Lunars were cheaper than Harrison's clock and were used extensively for the next century, eventually the clock became the standard for measuring longitude as it was accurate, simple, and more affordable. By the 1800s, chronometers were commonplace aboard English Navy and surveying ships; by 1815, over 5,000 marine chronometers were at sea. The problem of determining longitude at sea was finally solved.

Although compasses, almanacs, sextants, and clocks enabled worldwide naviga-

tion, position is not ascertainable with these methods in foggy or cloudy weather. Once the industrial age began, the position of large steel ships, sleek submarines, fast-moving aircraft and spacecraft had to be known more precisely. Technology's response was inertial instruments and radio navigation. The clock, however, remains an essential component for navigation.

1.3 Navigation in Modern Times

A gyroscope is one type of inertial instrument commonly used for navigation. A single-degree-of-freedom rate gyro used as a basic turn indicator for instrument flying first appeared in 1920 (*Barbour et al.*, 1992). A gyroscope employs a rotating mass whose rotational axis is in an inertial reference frame regardless of vehicle dynamics. Measuring the angle of rotation of the gyroscope enables accurate measurements of acceleration, which are integrated to obtain velocity and position. Many other inertial instruments work on similar principles, including accelerometers, fiber-optic gyros, ring-laser gyros, resonating beam accelerometers, microwave resonant accelerometers, and Microelectromechanical Systems (MEMS) (*Barbour et al.*, 1992). Collectively, a system of inertial instruments are known as an inertial measurement unit. These units require accurate initialization of the device's starting position, which is then propagated forward. Although inertial measurement units are very precise instruments, they are still prone to drifts over time. In the case of the mechanical gyroscope, drifts include imprecise knowledge of the varying gravity field of the earth or frictional forces. Often, updates to the position are required, and in some locations, such as a submarine in the middle of the ocean, precise position was not available until the advent of radio navigation.

Radio navigation is the technique of using radio waves to determine the position of a vessel. There are essentially three methods of radio navigation – trilateration (or time-of-arrival), hyperbolic positioning, and Doppler positioning.

Trilateration is a system that depends on the time of arrival of known radio signals from three or more transmitters to determine a 2-D position fix unambiguously. Trilateration requires that the transmitters and receiver clocks be synchronized, which is not an easily accomplished task. Fixed or moving radio transmitters such as GPS satellites can be used for trilateration.

Hyperbolic positioning is known as a time-difference-of-arrival system. For 2-D positioning, a hyperbolic positioning system requires three synchronized transmitting stations and a receiver that is capable of measuring only the difference in arrival times of the transmitted signals. There are various implementations of hyperbolic positioning, which include OMEGA and LORAN-C.

According to *Kasper and Hutchinson* (1978), OMEGA was the first true global (terrestrial) navigation system , and was decommissioned on 30 September 1997 after providing service for over 30 years. Worldwide accuracy was reported to be on the order of 1 nautical mile and was provided using only eight synchronized transmission stations and relatively inexpensive receivers (in 1978 an OMEGA receiver cost several thousand dollars). Phase synchronization was maintained with four cesium atomic frequency standards, and inter-station phase measurements were made twice daily to generate corrections.

LORAN-C aids air travel and ship traffic and has an accuracy of about 0.25 nm[1] over the continental United States. LORAN-C is currently still in use, and the susceptibility of GPS to jamming and space weather seems to have driven a

[1]LORAN-C General Information http://www.navcen.uscg.gov/loran/Default.htm

need for LORAN as a back up system.

Doppler positioning is a system whereby the received Doppler shift of a signal is recorded and then processed to determine the location of the receiver relative to the transmitter. A classic example of a system that used Doppler positioning is the Transit Navigation Satellite System, and was one of three enabling technologies that led to GPS.

1.4 Satellite Navigation and the Birth of GPS

The first satellite-based navigation system was the United States Navy Navigation Satellite System known as Transit. Developed at the Applied Physics Laboratory of The Johns Hopkins University, the first tests of the system were made in the 1960's. Transit consisted of 4 – 7 polar-orbiting satellites, at an altitude of 1100 km (*Parkinson*, 1996a). With Transit, a position fix could be obtained once every orbit (about 110 minutes) on the equator, and once every 30 minutes at 80° latitude. Transit was primarily used for marine navigation because a fix required 10 to 15 minutes of receiver processing in a low-dynamics environment with an estimate of the user's position.

By measuring the Doppler shift of the incoming signal from the satellite, the receiver could determine the user's position (*Stansell*, 1983). The shape of the Doppler shift curve allowed the user to calculate a distance from the satellite's ground track (known from the satellite ephemeris and the receiver clock). Transit only provided 2-D position fixes (latitude and longitude) and had a positioning accuracy of about 25 m (rms) for a stationary user. By tracking many satellite passes over successive days from a fixed location, a receiver could calculate a 3-D

position fix with 5 m accuracy. The main purpose of Transit was for inertial navigation system updates on the U.S. Navy Polaris strategic missile submarines.

The second enabling technology was the Navy satellite system known as Timation. Developed in 1972, Timation was developed primarily to distribute very precise time globally. Originally the satellites flew with quartz-crystal oscillators but later flew with rubidium and cesium atomic standards. The highly accurate clocks allowed for precise prediction of satellite ephemeris and less frequent intervention by the system operators (control segment). The Timation system pioneered and demonstrated that atomic standards could function onboard satellites and that global time synchronization was possible (*Parkinson*, 1996a).

The last enabling technology for GPS came from U.S. Air Force Project 621B. In 1972, this program demonstrated a new type of satellite-ranging signal based on pseudorandom noise signals (*Parkinson*, 1996a). A pseudorandom noise sequence is a repeated sequence of nearly random data-less bits known as chips that possess particularly useful properties. The sequences are easily generated using shift registers or can be stored in memory and read in sequence. By detecting the phase of the known pseudorandom noise sequence, a range to the satellite can be determined. Furthermore, a pseudorandom sequence can be detected in very weak signal environments, and inherently provides very good noise and interference rejection. Finally, pseudorandom noise sequences allow all satellites to broadcast on the same carrier frequency because properly selected pseudorandom noise sequences are orthogonal (the correlation of two differing pseudorandom sequences is nearly zero).

By the early 1970s the need for a global navigation and timing system that provided accurate 3-D position across all branches of the U.S. military became

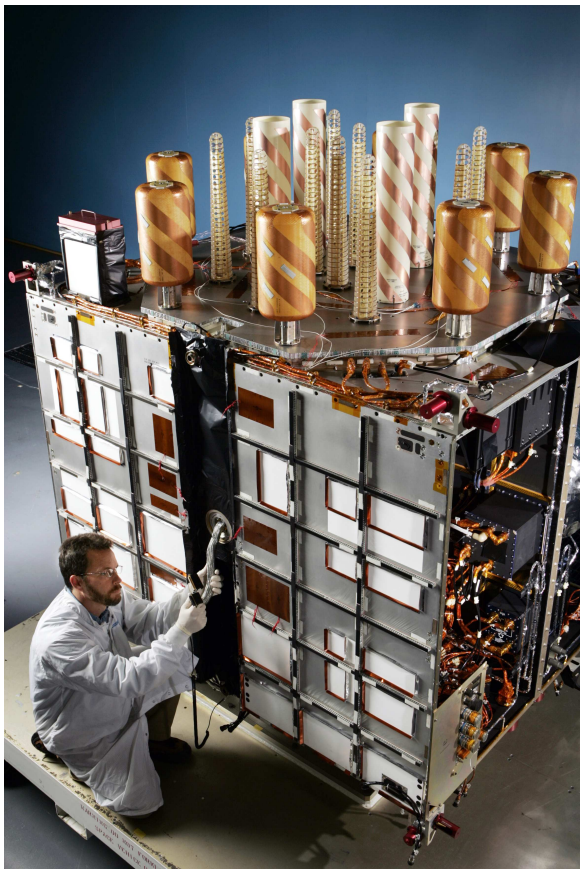

Figure 1.2: Photograph of a GPS Block II-RM satellite during construction. The transmitting antenna arrays are visible at the top. Image © Lockheed Martin Corporation, reproduced with permission.

apparent. Transit proved that satellite radio navigation was possible, but the sparsity of the constellation and the slowness of the Doppler positioning method was inadequate. Timation proved that atomic clocks and time synchronization was possible onboard satellites, and the U.S. Air Force project 621B demonstration proved that accurate range measurements to satellites was possible. The stage was set for a highly accurate 3-D satellite navigation system based on trilateration.

In 1973 the NAVSTAR GPS Joint Program Office (JPO) was created by the United States Department of Defense (DoD). The first director of the NAVSTAR GPS JPO was Dr. (Col.) Bradford W. Parkinson. The original name was given

by General Hank Stehling to be Global Positioning System, and NAVSTAR was tacked on to the name by a suggestion from Mr. John Walsh, who was a key decision maker for the GPS JPO budget. Contrary to popular belief, NAVSTAR is not an acronym for NAVigation System Timing and Ranging (*Parkinson*, 1996a).

Originally packaged as U.S. Air Force Project 621B, in August 1973 the Defense System Acquisition and Review Council rejected the first proposal for a satellite-based navigation system. Dr. Malcom Currie, Director of Defense Research and Engineering requested that the concept be expanded to include the views and requirements of all the armed services. Dr. Parkinson and twelve JPO members came together over Labor Day weekend of 1973 and hashed out the best of satellite navigation system concepts and technology. By 17 December 1973, approval for the program was granted.

Initially the program called for four demonstration satellites, and, by June 1974, Rockwell Collins had been selected as the satellite contractor. By February 1978, the first prototype satellite was launched, the control segment was deployed and working, and five types of user equipment (GPS receivers) were undergoing testing. At the end of 1994, GPS was declared to be "Full Operational Capability". Since then, the constellation has undergone significant improvements, upgrades, and replenishment, but the basic system design has been unchanged since the original concept was engineered in the mid-1970s. Finally, instantaneous and accurate position, velocity, and timing estimates were available on a global, continuous basis.

1.5 Foreign Satellite Navigation Systems

At this point it is worth mentioning a few words on other satellite navigation systems, which collectively are known as Global Navigation Satellite Systems (GNSS). The technological competition spurred by the Cold War between the Soviet Union and the United States drove each nation to match each other's inventions. As a result, the Soviets created satellite navigation systems of their own. The equivalent Soviet version of Transit was a system known as Tsikada (*Kaplan*, 1996), and their answer to GPS was the GLObal NAvigation Satellite System (GLONASS). GLONASS is still operational, and as of November 2007, the constellation consists of 9 operational satellites, 3 temporarily unavailable satellites, 4 satellites under decommissioning phase, 2 satellites to be commissioned, and 6 satellite slots that need to be filled[2]. Russian President Vladimir Putin has vowed to bring GLONASS back to full operational capability after it suffered from neglect in the post-Cold War era.

The Chinese are also developing a system known as Beidou. Currently the system has limited coverage and utilizes only three geostationary satellites. The system calls for 30 medium earth orbit (MEO) satellites and 5 geostationary orbit satellites. The 30 MEO satellites will operate in six orbital planes and will share many features in common with GPS to facilitate low-cost integration of the signals into GPS/GLONASS/Galileo/Beidou receivers. According to the International Telecommunication Union (ITU), Beidou will transmit at 1.58974 GHz, 1.5611 GHz, 1.26852 GHz, and 1.20714 GHz–very close to the existing GPS L1 (1.57542 GHz) and L2 (1.2276 GHz) signals. The system will provide an open civilian service and a higher precision encrypted military service. As of 14 April 2007,

[2]As indicated at http://www.glonass-ianc.rsa.ru

China launched the first Beidou satellite designated M-1 (*Gao et al.*, 2007).

Since 2003 the Europeans have been developing a system known as Galileo. Galileo is a public-private consortium that is funded in part by European tax payers and managed and further funded by private businesses. The goal of the system is to provide more accurate service than GPS or GLONASS, to improve positioning at high latitudes, and to provide an alternative, independent positioning system from GPS and GLONASS. The system is currently expected to be completed in 2011 or 2012, but the project continues to be plagued by hardware and political problems. In an 8 May 2007 article, the *Associated Press* reports that Galileo is in a deep crisis[3]. The companies set forth by the consortium have not agreed on how to divide the workload. European leaders suggested that the consortium is on the verge of collapse and will require further public funding. So far, the frequency allotments for Galileo have been reserved, and a satellite (known as Giove-A) launched in December 2005 has been transmitting Galileo-like codes and brief periods of navigation messages. The second test-bed satellite (Giove-B) was slated for launch in December 2006 but has been delayed indefinitely due to a catastrophic short-circuit during final testing[4]. Although the future of Galileo appears to be uncertain, one thing is certain: the advent of Galileo has spurred competition that has led to significant improvements to both GPS and GLONASS. Despite the problems with Galileo, in the end, most users will benefit from competing Global Navigation Satellite Systems, primarily due to increased precision from larger, more precise satellite constellations, either independently or as an aggregate of multiple systems.

[3]Associated Press, "EU: Galileo project in deep 'crisis'", http://www.cnn.com/2007/TECH/05/08/galileo.troubles.ap/index.html

[4]http://sidt.gpsworld.com/gpssidt/System+Design+and+Test+News/Galileo-Industries-Told-to-Put-House-in-Order/ArticleStandard/Article/detail/403494

Several augmentations to GNSS are currently in use or under development to address the immediate needs of GPS users who desire more precise services. A global effort exists to augment GPS using geostationary satellites that will provide single-frequency global ionospheric TEC and satellite clock and ephemeris corrections primarily for aviation use. In general, these systems are referred to as Satellite Based Augmentation Systems (SBAS). In the United States, this system is known as the Wide Area Augmentation System (WAAS), in Europe it is known as European Geostationary Navigation Overlay Service (EGNOS), in India it is known as GPS And GEO Augmented Navigation (GAGAN), and in Australia and Asia it is known as the Multi-functional Satellite Augmentation System (MSAS). The Japanese are also working on a system known as Quasi-Zenith Satellite System (QZSS) that employs geosynchronous satellites.

Although GPS has provided near real-time ultra-precise positioning capability, the augmentation systems were developed because errors still exist that are too large for some users to cope with. These errors include satellite clock and ephemeris prediction errors, ionospheric errors for single frequency users, and tropospheric errors. Other errors cannot be corrected using augmentation systems such as multipath and receiver dependent errors. Finally, all satellite navigation and augmentation systems are affected to some degree by space weather–either as loss of service, or as an increase in navigation errors.

1.6 Motivation for Understanding Space Weather Effects on GPS

In the GPS community, one often hears the saying 'One man's noise is another man's signal.' To the GPS engineers, space weather is a nuisance often considered to have a negligible effect on the performance of GPS receivers. This was especially true when Selective Availability[5] was still in effect and induced randomly varying errors of up to 100 meters. Space weather effects were inconsequential and difficult to differentiate from the purposeful degradation of the signal. GPS, however, has provided atmospheric scientists and space physicists with invaluable data on the upper atmosphere and near-space environment that has enabled significant advancement and discoveries in their respective fields. These scientists are currently busy developing models to predict space weather to minimize its impacts on modern society.

The termination of Selective Availability on 1 May 2000 brought a dramatic and instantaneous improvement on receiver positioning. To a lesser degree, the threat of competing Global Navigation Satellite Services spurred the United States Department of Defense to enhance the accuracy of the satellite ephemeris, which in turn improves the precision of the receivers. Most importantly, ultra-precise positioning and timing services for every purpose from time-stamping financial transactions to positioning oil rigs and navigating aircraft has increased the ante for the expected precision from GPS. While no where does the United States Government guarantee truly continuous and accurate (< 5 m) services with GPS[6], the

[5]Selective Availability was a purposeful degradation of the L1 coarse acquisition (publicly available) signal implemented by the Department of Defense.

[6]The *GPS SPS Performance Standard* (2001) requires at least 24 operational satellites be available 95% of the time over any day, with at least 21 transmitting satellites available 98% of the time in a year. The global average positioning error is ≤ 13 meters horizontally and ≤ 22

current capabilities of the system have launched business plans and expectations that assume the current level of service is, and will always be, available. Space weather effects on GPS are no longer negligible.

Although solar radio bursts were first discovered to affect radar systems during World War II, they were thought to have an unsubstantial effect on GPS receivers. This thesis presents the first quantitative analysis of the effect of solar radio bursts and builds upon the research conducted by *Klobuchar et al.* (1999) and *Chen et al.* (2005) through analysis of an event that occurred on 7 September 2005. Although the event was quite modest, it quantitatively verified the suggestion by *Chen et al.* (2005) that the threshold that was deemed deleterious to receiver operation by *Klobuchar et al.* (1999) had to be revisited. Given the sensitivity of GPS receivers, *Cerruti et al.* (2006) suggested that receivers could be vulnerable to large solar radio bursts. The truth occurred sooner than expected, and in December 2006 a series of solar radio bursts seriously challenged, and in some cases completely caused loss of navigation on sunlit receivers.

1.7 Thesis Structure

This thesis is divided as follows. In Chapter 2, space weather as it applies to GPS will be analyzed. Space weather primarily challenges the receiver's ability to accurately determine its position. The effects could be as mundane as the ever-present delay caused by the ionosphere, or as severe as scintillations caused by plasma irregularities associated with convective ionospheric storms. The response of the ionosphere to solar flares and geomagnetic storms will also be analyzed.

meters vertically, and at most ≤ 36 meters horizontally and ≤ 77 meters vertically. Time transfer is guaranteed to be better than 40 ns.

Solar radio bursts will be introduced, only to be explored in later chapters.

Chapter 3 is designed to give the reader background information on GPS. GPS operations and performance requirements will be detailed, and in particular, the GPS signal structure, navigation message, and signal power will be detailed as their understanding will motivate the vulnerability of GPS to solar radio bursts, discussed in Chapter 4. Finally, GPS receivers will be discussed at a high-level, primarily to introduce the signal processing techniques employed within the receivers. This discussion will help to motivate solar radio burst mitigation techniques, discussed in Chapter 6.

Chapter 4 presents a computational analysis that will show the expected effect of a solar radio burst's impact on a GPS receiver. The impact on coded receivers as well as semi-codeless receivers (for tracking the encrypted military code) will be analyzed. Furthermore, the impact of decreased carrier-to-noise ratio and its effects on the receiver tracking loops and measurement variances will be analyzed. The computational analysis will be verified using a moderate solar radio burst that impacted GPS on 7 September 2005. As part of the analysis of the day's events, other potential sources of space weather degradation are ruled out. The chapter concludes with new predictions for the threshold at which solar radio bursts should impact receivers.

The predictions made in Chapter 4 came sooner than predicted. In December 2006, three historical solar radio bursts occurred, and are all the more surprising because they occurred at solar minimum. The largest solar radio burst, which occurred on 6 December 2006, severely impacted the ability of receivers in the sunlit hemisphere to adequately track the satellite signals. In many cases, receivers failed to produce a navigation solution. Those receivers that did not fail to produce

a navigation solution saw drastic increases in positioning errors due to difficulty in tracking the weak GPS signals. The performance of several GPS services including the Federal Aviation Administration's (FAA) Wide-Area Augmentation System (WAAS) and the International GNSS Service are detailed.

Although the December 2006 solar radio bursts severely challenged a GPS receiver's ability to track the satellite signal, there exist several ways to mitigate the effects of a GPS receiver to a solar radio burst. Techniques such as narrowing the tracking loop bandwidth, or externally aiding the receiver with inertial navigation systems significantly improve the receiver's ability track through strong interference. Moreover, additional planned signals and additional navigation satellites from competing systems will help to mitigate the effects of solar radio bursts. These techniques, and several others, are explored in Chapter 6. In this chapter, a technique to determine solar radio burst flux estimates from GPS receivers is developed. The December 2006 events suggest that the solar radio burst record may be inaccurate due to saturation of dedicated monitoring hardware. The technique to determine solar radio burst flux developed in this chapter could prove useful because it will allow past solar radio burst events to be re-analyzed. The thesis concludes with remarks on the broader impacts of the solar radio burst, including societal impacts.

Chapter 2

Space Weather

A great, great deal has been said about the weather, but very little has ever been done. - Mark Twain

2.1 Introduction

The sun is ultimately the energy source for all space weather here on Earth. In the universe, the sun is a rather average star with a mass of $1.99 \cdot 10^{30}$ kg, a radius of 696,000 km, and a luminosity of $3.86 \cdot 10^{26}$ W. It is composed of approximately 90% hydrogen, and about 10% helium, and about 0.1% other elements. The sun is almost entirely plasma, with a fusion reactor operating at temperatures of 16 million kelvin at its core as the energy source. Every second, over 4 million metric tons of matter (mostly hydrogen) is converted to energy in a careful balance between outward radiative pressure and inward gravitational pressure. The energy generated by the fusion processes slowly radiates out through the radiative zone and then in to the convective zone, which begins about 500,000 km from the center.

The surface of the sun appears granulated when observed through a solar telescope and has a temperature of roughly 5,800 Kelvin. Strong convection rising through the convective zone is responsible for the visible structuring. Bright areas are representative of heat escaping from within, and dark, cooler areas mark the boundaries of the convective cells. The granules range in size from 1,000 km - 30,000 km and last from 10 minutes to as long as an Earth day.

The sun's atmosphere is divided into three regions–the photosphere, chromo-

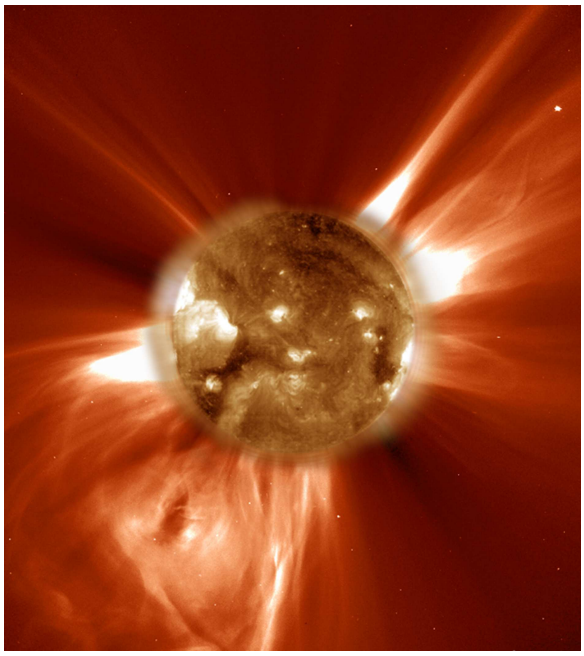

Figure 2.1: Composite image of the sun in ultraviolet light created by the Solar and Heliospheric Observatory (SOHO), showing a coronal mass ejection (bottom left), coronal streamers (top right), and coronal holes (dark areas on the surface). Image courtesy of NASA/JPL.

sphere, and corona. The photosphere is a thin, relatively cool layer (∼5,000 K) just above the convective zone from which all the visible light is emitted. The thin chromosphere lies just above the photosphere and is characterized by a rapidly rising temperature from a minimum of 4,500 K at the interface with the photosphere, to about 25,000 K at the base of the chromosphere. The corona is the highest region of the sun's atmosphere and extends out to roughly 10 solar radii. It contains a very tenuous, 1 million kelvin plasma that is thought to be heated by magnetic reconnection.

Close to the sun, the solar magnetic field lines close back to the surface, trapping the plasma. In the corona, however, the high temperatures cause a continual shedding of plasma, known as the solar wind. The solar wind drags the sun's magnetic structure into interplanetary space. Coronal plasma is trapped in structures

known as coronal loops, and weaker forms of these structures, known as coronal streamers, form the slow component of the solar wind. Some magnetic field lines do not loop back into the sun but rather extend radially out into space. Plasma from the surface of the sun can readily escape along the radial magnetic field lines forming the fast solar wind. These areas are known as coronal holes.

The sun does not rotate as a rigid body. The poles rotate more slowly at 34 days, and the equator rotates with a period of roughly 25 days. The differential rotation, coupled with plasma convection from the hot core, produce intense electric currents and magnetic fields through dynamo action. The intensity of the magnetic fields wax and wane with a period of 11 years, which corresponds to a reversal of the sun's magnetic polarity. The changing magnetic field lines become tangled and twisted together and can trap plasma and form structures known as filaments and prominences. Sometimes the magnetic fields become so twisted that they prevent hot plasma from convecting to the surface. These regions, known as sunspots, appear as dark regions with diameters up to 20,000 km and are regions of relatively cooler plasma (4,000 - 4,500 kelvin versus 5,800 kelvin). Sunspots have lifetimes of hours to weeks.

The number, frequency, and location of sunspots change with a period of roughly 11 years, known as the solar cycle. At the beginning of the solar cycle, the sunspots tend to be located higher in latitude, and as the cycle progresses they move equatorward. The number of sunspots maximizes at solar maximum, after which the sun's magnetic field is destroyed. This period of activity typically lasts from 3-5 years. Through magnetic reconnection, the magnetic fields above sunspots may reconfigure and unleash powerful explosions of visible, ultraviolet, and x-ray light known as solar flares. Solar flares also eject high-energy particles

which can interact with the surrounding plasma and can create solar radio bursts, which are large and sudden increases of radio energy (from a few MHz to tens of GHz) as compared to background radio emissions from the sun.

Particle ejection also occurs when one end of the magnetic field containing a prominence breaks away from the surface of the sun. Even more dramatically, coronal mass ejections (CMEs) are massive explosions of matter from the sun's corona that tend to coincide with solar flares. Once again, magnetic reconnection provides the energy for these massive explosions. Coronal mass ejections eject as much as 10^{13} kg of plasma with velocities that can exceed 1,000 km/s.

Solar wind shocks created by the transition from fast to slow-moving solar wind regions as well as those created by CMEs arrive as super-sonic shocks that buffet the Earth's magnetosphere. Often the shocks also carry high energy particles that penetrate into the polar regions along magnetic field lines that connect with the solar wind magnetic field. Under the appropriate conditions, these shocks induce global-wide magnetic storms. During a geomagnetic storm, the Earth's magnetosphere is altered by the fast-moving solar wind shock. Magnetic reconnection induced by the shockwave highly accelerates electrons and protons that stream into the polar regions, inducing the aurora. The Van Allen radiation belts are rearranged, resulting in a ring current that creates a magnetic field opposite to the Earth's magnetic field at the surface of the Earth. The rearranging of the magnetosphere as well as the induced electric and magnetic fields from the Van Allen radiation belts drive global-wide disturbances of the ionosphere, known as ionospheric storms.

Much as Mark Twain said that "A great, great deal has been said about the weather, but very little has ever been done", the complete theory, experiments,

and observations of space weather and plasma physics is an immensely broad subject and many books have been written (e.g. *Schunk and Nagy* (2000b), *Kivelson and Russell* (1995), *Kelley* (1989), and *Davies* (1965)). Our ability to predict and model space weather is still an active, and growing area of research that is vital to ensure our safety during an increasingly technologically-dependent society. Currently, the United States National Oceanographic and Atmospheric Administration's (NOAA) National Weather Service (NWS) Space Weather Prediction Center[1] is the primary source of space weather forecasts.

In the following sections, space weather and its impacts on GPS will be considered, primarily to understand their effect on receiver operation. Therefore, the physical processes that generate the space weather will only be explained to the extent that they aid the discussion.

With one important exception, the ionosphere and its myriad of dynamical processes are responsible for space weather effects on the Global Positioning System (GPS). The Earth's ionosphere is a plasma created by the ionization of neutral gas by solar x-ray and ultraviolet radiation. The ionosphere is globally present and causes dispersion of radio signals propagating through it. Discussed in Section 2.2, the dispersion of the signal induces a delay in measuring the distances to the satellites, which in turn affects the navigation accuracy. Solar flares directly and rapidly increase the ionization in the sunlit hemisphere, resulting in further ranging errors, and are discussed in Section 2.3.

Other space weather effects on GPS can be categorized by geomagnetic longitude. In equatorial regions (Section 2.4), GPS is principally affected by the high ionospheric plasma density, which is enhanced by the equatorial ionization

[1] http://www.swpc.noaa.gov/

anomalies. The night-time equatorial ionosphere is also prone to plasma irregularities known as convective ionospheric storms (manifested as equatorial spread-F on radar returns and ionosondes) that cause large amplitude and phase scintillations that challenge GPS satellite signal tracking.

The mid-latitude ionosphere is principally dominated by geomagnetic storms, both of which are discussed in Section 2.5. Geomagnetic storms increase the radiation hazard to the GPS satellites, which are in orbits that place them within the Van Allen radiation belts. Geomagnetic storms also restructure the ionosphere and sometimes convect poleward plumes of dense equatorial plasma known as storm-enhanced density (SED) (*Foster and Rideout*, 2005). Large TEC gradients and mid-latitude scintillation have been observed during SED events (e.g. *Basu et al.* (2005b) and *Ledvina et al.* (2002)). The equatorial ionization anomaly are driven upward and poleward (*Foster and Rich*, 1997). Geomagnetic storms are also responsible for equatorward advances of the auroral oval.

At high latitudes, the aurora and ionospheric irregularities again affect the GPS signal. Although these phenomenon can occur frequently, they rarely have a significant impact on GPS signals, but are shown nonetheless in Section 2.6.

The one space weather exception that is not related to the ionosphere that affects the Global Positioning System are solar radio bursts. Solar radio bursts of the appropriate frequency affect all sunlit GPS receivers by acting as a wide-band interference source. Solar radio bursts will be briefly introduced in Section 2.7 since their impact on GPS is the topic of this thesis.

2.2 The Ionosphere

The Earth's ionosphere is a region of plasma that extends from roughly 100 km to 1000 km in altitude. Solar ultraviolet and x-ray radiation is predominantly responsible for ionizing the gasses of the upper atmosphere. The solar ionization is in constant competition with chemical recombination. Solar ionization is a function of the solar cycle (more ionization during solar maximum), the season, and the solar elevation angle. Recombination rates vary primarily as a function of atmospheric density, and hence altitude. The electron density is also affected by neutral winds and gravity waves.

Vertically, the ionosphere is primarily organized into layers or regions. These regions are determined by the chemical composition of the atmosphere. The upper atmosphere is stratified predominantly by molecular and atomic gasses that have different ionization potentials. Although it is well beyond the scope of this thesis to discuss the various photo-ionization processes, the layers of the ionosphere will be briefly discussed.

The lowest layer of the ionosphere is known as the D-region and extends from roughly 50 - 90 km. The D-region is ionized by hard x-ray wavelengths, but remains weakly ionized during the day because the atmospheric density is relatively high, and hence recombination is also high. The D-region becomes more ionized during solar flares and the increased ionization, in combination with the dense neutral atmosphere, can lead to HF radio blackouts.

The next layer is known as the E-region and extends from roughly 90 km - 120 km and is ionized predominantly by soft x-rays and far ultraviolet solar radiation. The E-region is less influenced by recombination because atmospheric

densities are lower, so it has higher ionization levels than the D-region. During the day, the E-region is highly conductive and the electric fields control the overall structure of the ionosphere. At night, however, the E-region rapidly recombines and effectively becomes an insulator. At night, then, the F-region electric fields dominate the overall structure.

The uppermost layer of the ionosphere is known as the F-region. The peak electron density of the ionosphere occurs in the F-region, where recombination rates are low, and day-time photo-ionization by ultraviolet light is high. The peak electron density is variable but is often at an altitude of roughly 350 km, with electron densities on the order of 10^{12} e^-/m^3. Above 350 km, the neutral gas species rarify, so the electron density of the ionosphere decreases steadily until it merges with the plasmasphere. Although the F-region densities typically decrease at night, the gasses are so rarefied at these heights that recombination is slow and the F-region ionosphere often persists until sunrise. Total peak ionization for all layers occurs in the local afternoon hours, with an abrupt end of ionization after local sunset. At local sunrise, the ionization immediately rises again. Figure 2.2 shows a modelled vertical electron density profile over Arecibo Observatory (18.34°N, 66.754°W) during a geomagnetically quiet period ($\langle K_p \rangle \approx 1.14$) at solar minimum.

2.2.1 Ionospheric Dispersion

The ionosphere is a magnetized plasma in which a multitude of interesting wave dispersion relations exist. In a dispersive medium, the phase and group velocities of a wave propagating through it are functions of the wave frequency and wave vector direction. GPS uses radio waves from trans-ionospheric trilateration, so

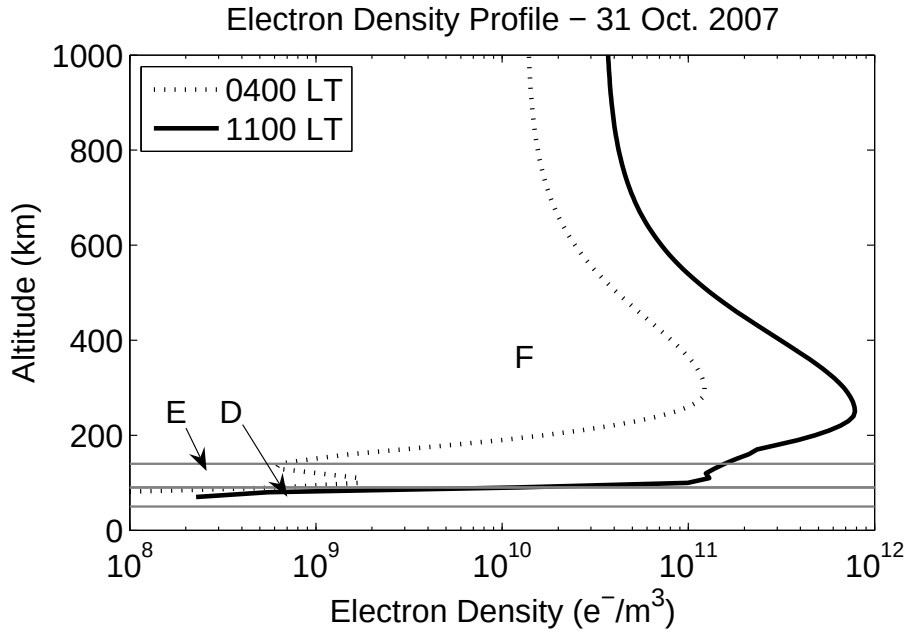

Figure 2.2: Vertical ionospheric profile densities as derived from the International Reference Ionosphere model over Arecibo Observatory (18.34°N, 66.754°W) at solar minimum ($\langle K_p \rangle \approx 1.14$). Notice the distinct differences in the ionosphere between night and day and that the F-layer never entirely disappears at night. At 1100 LT, the ionosphere over Arecibo was nearly at its highest density for the day. As an indication of the model's accuracy, the vertical TEC as derived from the model is 19.3 TECU, whereas the TEC derived from GPS instruments is 18.8 TECU. IRI model data from SPDF/Modelweb (http://modelweb.gsfc.nasa.gov/models/iri.html). K_p data recovered from Space Physics Interactive Data Resource.

the signal propagation through the dispersive ionosphere has important effects on measuring the range to the satellite that must be considered.

We begin first by reviewing the derivation of an electromagnetic wave in a vacuum. The following derivations are based on *Chen* (1984). Suppose the electromangetic wave from the GPS satellites is as defined in Figure 2.3, where $\mathbf{E}_1 = E_1\hat{\mathbf{z}}$ and $\mathbf{B}_1 = -B_1\hat{\mathbf{y}}$ refer to the electric and magnetic fields of the wave.

Maxwell's Equations in a vacuum are:

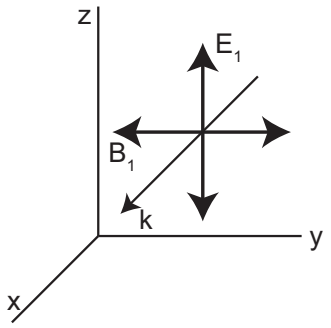

Figure 2.3: Geometry for electromagnetic wave propagation with $\mathbf{k} \perp \mathbf{B_0}$.

$$-\dot{\mathbf{B}}_1 = \nabla \times \mathbf{E}_1 \tag{2.1}$$

$$\dot{\mathbf{E}}_1 = c^2 \nabla \times \mathbf{B}_1 \tag{2.2}$$

$$\nabla \cdot \mathbf{E}_1 = 0 \tag{2.3}$$

$$\nabla \cdot \mathbf{B}_1 = 0 \tag{2.4}$$

where $c = 1/\mu_0\epsilon_0$ is the speed of light in a vacuum. Taking the curl of Equation 2.2 yields:

$$\nabla \times \dot{\mathbf{E}}_1 = c^2 \nabla \times (\nabla \times \mathbf{B_1}) \tag{2.5}$$

which is substituted into Equation 2.1:

$$-\ddot{\mathbf{B}}_1 = c^2 \nabla \times (\nabla \times \mathbf{B_1}) \tag{2.6}$$

We further assume that the wave is planar and therefore has a form $e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, which allows the derivatives and curls in Equation 2.6 to be evaluated[2]:

[2] $\mathbf{k}$ is the wave vector. $k = \frac{2\pi}{\lambda}$, where λ is the wavelength of the electromagnetic wave

$$\begin{aligned} \omega^2 \mathbf{B}_1 &= -c^2 \mathbf{k} \times (\mathbf{k} \times \mathbf{B}_1) \\ &= -c^2 \left[\mathbf{k} (\mathbf{k} \cdot \mathbf{B}_1) - k^2 \mathbf{B}_1 \right] \end{aligned} \tag{2.7}$$

$(\mathbf{k} \cdot \mathbf{B}_1) = 0$ by Equation 2.4, resulting in:

$$\omega^2 = k^2 c^2 \ (\text{rad/s})^2 \tag{2.8}$$

The phase velocity, v_ϕ and the group velocity v_g can be derived from Equation 2.8 and are:

$$v_\phi = \frac{\omega}{k} = c \text{ m/s} \tag{2.9}$$

$$v_g = \frac{d\omega}{dk} = c \text{ m/s} \tag{2.10}$$

As expected, the phase and group velocities of an electromagnetic wave in a vacuum are equal to the speed of light, c.

When the GPS signal enters the ionosphere, it encounters a plasma. The oscillating electromagnetic wave induces currents in the plasma that cause the electromagnetic wave to be mechanically damped. The interaction of the wave with the surrounding plasma yields many dispersion relations for the phase and group velocities of the wave that depend on the direction of the magnetic field with respect to the wave vector, $\mathbf{k}$.

The transmitted GPS signal propagates transverse to the background magnetic field embedded with the plasma, but for purposes of this discussion, the magnetic

field will be neglected. In a plasma, the Maxwell Equations (Equations 2.2 - 2.4) become:

$$-\dot{\mathbf{B}}_1 = \nabla \times \mathbf{E}_1 \tag{2.11}$$

$$\frac{\mathbf{j_1}}{\epsilon_0}\dot{\mathbf{E}}_\mathbf{1} = c^2 \nabla \times \mathbf{B_1} \tag{2.12}$$

$$\nabla \cdot \mathbf{E}_1 = 0 \tag{2.13}$$

$$\nabla \cdot \mathbf{B}_1 = 0 \tag{2.14}$$

where $\mathbf{j}_1/\epsilon_0$ represents the currents due to first-order charged particle motions.

We proceed again by taking the curl of Equation 2.11:

$$\begin{aligned} -\nabla \times \dot{\mathbf{B}}_1 &= \nabla \times (\nabla \times \mathbf{E_1}) \\ &= \nabla (\nabla \cdot \mathbf{E}_1) - \nabla^2 \mathbf{E}_1 \end{aligned} \tag{2.15}$$

Next, the derivative of Equation 2.12 is:

$$\frac{1}{\epsilon_0}\frac{\partial \mathbf{j}_1}{\partial t} + \ddot{\mathbf{E}}_1 = c^2 \nabla \times \dot{\mathbf{B}}_1 \tag{2.16}$$

which is substituted into Equation 2.15:

$$\frac{-1}{\epsilon_0 c^2}\frac{\partial \mathbf{j}_1}{\partial t} - \ddot{\mathbf{E}}_1 = \nabla (\nabla \cdot \mathbf{E}_1) - \nabla^2 \mathbf{E}_1 \tag{2.17}$$

that is now only dependent on $\mathbf{E}_1$. Again, assuming a planar wave dependence of $e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, and setting $\nabla = i\mathbf{k}$ and $\frac{\partial}{\partial t} = -i\omega$ we obtain:

$$\frac{i\omega}{\epsilon_0 c^2}\mathbf{j}_1 + \frac{\omega^2}{c^2}\mathbf{E}_1 = -\mathbf{k}\left(\mathbf{k}\cdot\mathbf{E}_1\right) + k^2\mathbf{E}_1 \tag{2.18}$$

The electromagnetic waves are transverse so $\mathbf{k}\cdot\mathbf{E}_1 = 0$ and Equation 2.18 becomes:

$$\begin{aligned} \frac{i\omega}{\epsilon_0 c^2}\mathbf{j}_1 + \frac{\omega^2}{c^2}\mathbf{E}_1 &= k^2\mathbf{E}_1 \\ \frac{-i\omega\mathbf{j}_1}{\epsilon_0} &= \left(\omega^2 - c^2k^2\right)\mathbf{E}_1 \end{aligned} \tag{2.19}$$

We assume that only the electrons are mobil enough to respond to the wave and the heavy ions remain unaffected. The $\mathbf{j}_1$ current, therefore, is due entirely to the motion of the electrons:

$$\mathbf{j}_1 = -n_0 e\mathbf{v}_e \text{ A/m}^2 \tag{2.20}$$

where n_0 is the electron density of the plasma (in electrons per cubic meter - e^-/m^3), $e = 1.60217646\cdot10^{-19}$ C is the charge of an electron, and $\mathbf{v}_e$ is the velocity of the electrons. The electrons will be accelerated by an electric field according to Newton's Second Law of Motion:

$$m\frac{\partial\mathbf{v}_e}{\partial t} = -e\mathbf{E} \text{ N} \tag{2.21}$$

where $m = 9.1093897 \cdot 10^{-19}$ kg is the electron rest mass. The electric field that sets the electron into motion is from the electric field of the electromagnetic wave, which allows us to set $\mathbf{E} = \mathbf{E_1}$. Again with planar wave dependence of $\mathrm{e}^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, then:

$$\mathbf{v}_e = \frac{eE_1}{im\omega} \text{ m/s} \tag{2.22}$$

Substituting Equations 2.22 and 2.20 into Equation 2.19 yields:

$$\frac{i\omega}{\epsilon_0} n_0 e \frac{e\mathbf{E_1}}{im\omega} = \frac{n_0 e^2}{\epsilon_0 m}\mathbf{E_1} = \left(\omega^2 - c^2k^2\right)\mathbf{E_1} \tag{2.23}$$

$\sqrt{\frac{n_0 e^2}{\epsilon_0 m}} = \omega_p$ is defined as the plasma frequency, and $\epsilon_0 = 8.854 \cdot 10^{-12}$ F/m is the permittivity of free space. $\mathbf{E_1}$ cancels and we obtain the following dispersion:

$$\omega^2 = \omega_p^2 + c^2k^2 \text{ (rad/s)}^2 \tag{2.24}$$

Solving for v_ϕ and v_g:

$$v_\phi = \frac{\omega}{k} = \frac{c}{\sqrt{1 - \frac{\omega_p^2}{\omega^2}}} \text{ m/s} \tag{2.25}$$

$$v_g = \frac{d\omega}{dk} = c\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \text{ m/s} \tag{2.26}$$

Equation 2.25 implies that the phase velocity of the electromagnetic wave in the ionosphere is always greater than the speed of light in a vacuum. Conversely, the

group velocity is less than the speed of light. There are no violations of Einstein's Theory of Relativity because v_g never exceeds the speed of light.

When the electromagnetic wave has a frequency below that of the plasma frequency, w_p, the electromagnetic wave cannot propagate through the plasma and is reflected. In this case the radical in Equations 2.25 and 2.26 evaluates to an imaginary number, and the wave is cutoff. Cutoff typically occurs at frequencies around $f_p = \omega_p/2\pi = 10$ MHz and rarely exceeds 30 MHz in the ionosphere. Plasma densities in the ionosphere are never severe enough to cause cutoff for GPS since the transmitted L1 (1.57542 GHz) and L2 (1.2276 GHz) frequencies are so high. A physically impossible electron density of $1.87 \cdot 10^{16}$ e^-/m^3 is required to cause cutoff at L2, and the density is even higher at L1.

2.2.2 Ranging Error and TEC

The only parameter that varies as a function of space in Equations 2.25 and 2.26 is the electron density, n_0. The delay, δt, induced by the ionosphere causes an advance of the carrier phase that is equal to the delay of the wave envelope. The delay on the wave envelope, or, in the context of GPS, the pseudorange, is determined from the group velocity, v_g, and can easily be calculated.

The delay, δt, is the additional time the signal takes to travel through a length of plasma, ρ at velocity v_g, as opposed to an equal length traveled through a vacuum at velocity c, is:

$$\delta t = \int_\rho \left[\frac{1}{v_g} - \frac{1}{c}\right] d\rho \text{ s} \tag{2.27}$$

If we consider that the GPS L1 (1.57542 GHz) and L2 (1.2276 GHz) frequencies are much larger than the highest plasma frequency (at most 30 MHz), then $\frac{1}{v_g}$ can be approximated as:

$$\frac{1}{v_g} = \frac{1}{c\sqrt{1-\frac{\omega_p^2}{\omega^2}}} \simeq \frac{1}{c}\left(1+\frac{1}{2}\frac{\omega_p^2}{\omega^2}\right) \text{ s/m} \tag{2.28}$$

so then Equation 2.27 evaluates to:

$$\begin{aligned} \delta t &= \int_\rho \left[\frac{1}{c}\left(1+\frac{1}{2}\frac{\omega_p^2}{\omega^2}\right)-\frac{1}{c}\right] d\rho \Big|_{w_p^2=\frac{n_0(\rho)e^2}{\epsilon_0 m}, w^2=(2\pi f)^2} \\ &= \frac{e^2}{2c\epsilon_0 m(2\pi)^2 f^2}\int_\rho n_0(\rho)d\rho \text{ s} \end{aligned} \tag{2.29}$$

where $n_0(\rho)$ are variations in electron density along the signal path, ρ. The quantity $\int_\rho n_0(\rho)d\rho$, when evaluated, is known as the Total Electron Content or TEC, and is representative of the number of free electrons in a one-square-meter column along the signal path from the receiver to the satellite. Evaluating the constants in front of the integral and substituting $\int_\rho n_0(\rho)d\rho = TEC$ yields:

$$\delta t \simeq \frac{40.3}{cf^2} TEC \text{ s} \tag{2.30}$$

These equations do not account for dispersion caused by the ordinary nor extraordinary mode of propagation where the Lorentz force on the electrons must be taken into account. For most GPS applications, where accuracy to the meter level is acceptable, these terms are negligible. The second-order and third-order effects are on the order of $\sim 0-2$ cm, and $\sim 0-2$ mm at zenith (*Bassiri and Hajj*,

1993). For further information on these effects, please refer to *Hoque and Jakowski* (2007) and *Bassiri and Hajj* (1993).

TEC, in general, is a rather large quantity, with values on the order of $10^{16} - 10^{18}$ e^-/m^2. So, a unit known as a TEC Unit (TECU) is generally adopted, where 1 TECU = 10^{16} e^-/m^2. For GPS L1, each TECU contributes a delay of $\delta t = 0.5417$ ns, or an equivalent ranging error of 0.1624 m.

With only a single frequency, it is not possible to measure the ionospheric delay induced on the signal. Dual-frequency GPS receivers, however, are capable of tracking both the L1 and L2 signals simultaneously, allowing the receiver to calculate the ionospheric delay.

The delay in arrival time $\Delta(\delta t)$ between the L2 and L1 signals is calculated as:

$$\begin{aligned} \Delta(\delta_t) &= \frac{40.3}{cf_{L2}^2} TEC - \frac{40.3}{cf_{L1}^2} \text{ s} \\ &= \frac{40.3}{c} \left(\frac{f_{L1}^2 - f_{L2}^2}{f_{L1}^2 f_{L2}^2} \right) TEC \text{ s} \end{aligned} \tag{2.31}$$

where $f_{L1} = 1.57542$ GHz and $f_{L2} = 1.2276$ GHz. The difference in arrival times from L1 to L2 for one TECU is $\Delta(\delta_t) = 0.3504$ ns, which corresponds to 0.1050 m.

Determining the TEC between the satellite and the receiver is not as trivial as the derivations above. In reality, differential path delays within the circuitry of the satellite and receiver add biases between the L1 and L2 signals which are indecipherable from the measured ionospheric TEC. These differential biases, known as satellite and receiver TEC biases, must be accounted for when determining TEC to the satellite, and the process is not trivial.

Several services including the International GNSS Service (IGS) and NASA Jet Propulsion Laboratory (JPL) calculate and provide the TEC biases for the satellites and for a handful of ground-based geodetic receivers using extensive algorithms[3]. The differential satellite bias is also uploaded to each satellite and retransmitted as the T_{GD} parameter of the navigation message. Determining the receiver TEC bias is complicated and requires extensive calculation. *Rideout and Coster* (2006) offer a few explanations for determining receiver TEC bias.

In general, however, there are two approaches to determine receiver TEC bias. The first method minimizes the scalloping of the TEC curves once the obliquity factor has been removed. The obliquity factor accounts for the additional ionization encountered by a low-elevation signal as it propagates through the ionosphere. Another common method is to estimate the TEC attributable to the ionosphere by using models. These models use large networks of receivers to accurately estimate (2 - 8 TECU residual error) the state of the ionosphere. The TEC values from the model are subtracted from the TEC values determined by the receiver yielding a measure of the receiver TEC bias. One such model is the Global Ionospheric Map (GIM) that assimilates data from over 100 GPS receivers to produce global TEC maps (*Mannuci et al.*, 1998). The data is available in the IONEX files published by JPL/NASA through IGS.

In general, at solar minimum, for most latitudes, the ionosphere contributes roughly 5 - 10 TECU when observing a satellite directly overhead, as shown in the top panel of Figure 2.4. At solar maximum, the vertical delay associated with the ionosphere can exceed 150 TECU, which is a ranging error of over 30 meters at GPS L1. The bottom panel of Figure 2.4 illustrates the ionospheric

[3]The differential biases are available at http://igscb.jpl.nasa.gov/components/prods.html and are contained within the IONEX files which are found by following the 'Ionospheric TEC grid' links.

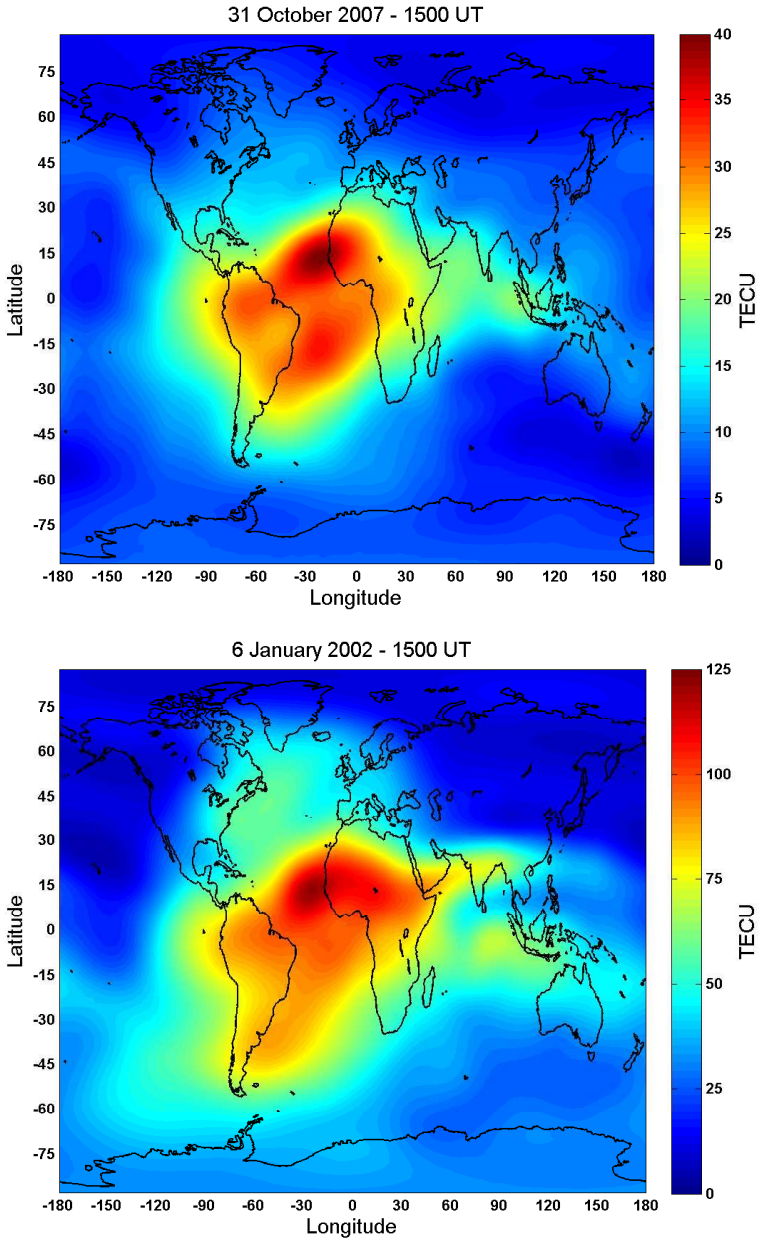

Figure 2.4: The top panel shows the vertical Total Electron Content of the ionosphere at solar minimum (geomagnetically quiet $\langle K_p \rangle < 1.14$) created from worldwide dual-frequency observations of GPS satellites. The bottom panel illustrates the vertical Total Electron Content of the ionosphere at solar maximum (geomagnetically quit $\langle K_p \rangle \approx 0.48$. The color scales are *not* equal. Notice the large differences between the ionosphere between solar maximum and solar minimum. World TEC map data recovered from JPL IONEX files (available from the International GNSS Service website (*Dow et al.*, 2005)). K_p data recovered from Space Physics Interactive Data Resource.

density near solar maximum. Note the large-scale differences from solar minimum to solar maximum. Considering that the GPS error budget is typically around 10 - 15 meters for single-frequency users, 30 meters of ranging error to each satellite would be a significant degradation to the service.

2.2.3 Correcting for the Ionosphere

Dual frequency pseudorange measurements, P_{L1} and P_{L2}, made at L1 and L2 respectively, allow an 'ionosphere-free' pseudorange, P_{IF}, to be determined. From Equation 2.31, multiplying both sides by c, we can relate the delay $c\Delta(\delta_t)$ to the difference of the measured pseudoranges at the two frequencies $(P_{L2} - P_{L1})$:

$$\begin{aligned} c\Delta(\delta_t) &= 40.3 \left(\frac{f_{L1}^2 - f_{L2}^2}{f_{L1}^2 f_{L2}^2} \right) TEC \text{ m} \\ P_{L2} - P_{L1} &= 40.3 \left(\frac{f_{L1}^2 - f_{L2}^2}{f_{L1}^2 f_{L2}^2} \right) TEC \text{ m} \end{aligned} \tag{2.32}$$

The TEC is determined by rearranging Equation 2.32 to obtain:

$$TEC = \frac{P_{L2} - P_{L1}}{40.3} \left(\frac{f_{L1}^2 f_{L2}^2}{f_{L1}^2 - f_{L2}^2} \right) \text{ e}^-/\text{m}^2 \tag{2.33}$$

A measure of the ionosphere-free pseudorange (P_{IF}) can be made by taking the difference of the L2 pseudorange and the estimate of TEC from Equation 2.33:

$$
\begin{aligned}
P_{IF} &= P_{L2} - \frac{40.3}{f_{L2}^2} TEC \\
&= P_{L2} - \frac{40.3}{f_{L2}^2} \frac{P_{L2} - P_{L1}}{40.3} \left(\frac{f_{L1}^2 f_{L2}^2}{f_{L1}^2 - f_{L2}^2} \right) \\
&= P_{L2} - \frac{P_{L2} - P_{L1}}{f_{L2}^2} \left(\frac{f_{L1}^2 f_{L2}^2}{f_{L1}^2 - f_{L2}^2} \right) \\
&= P_{L2} - (P_{L2} - P_{L1}) \left(\frac{f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \right) \\
&= P_{L2} - \frac{P_{L2} f_{L1}^2}{f_{L1}^2 - f_{L2}^2} + \frac{P_{L1} f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \\
&= \frac{P_{L2} f_{L1}^2 - P_{L2} f_{L2}^2}{f_{L1}^2 - f_{L2}^2} - \frac{P_{L2} f_{L1}^2}{f_{L1}^2 - f_{L2}^2} + \frac{P_{L1} f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \\
&= \frac{-P_{L2} f_{L2}^2 + P_{L1} f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \\
&= \frac{-P_{L2} + \gamma P_{L1}}{\gamma - 1} \\
&= \frac{P_{L2} - \gamma P_{L1}}{1 - \gamma} \text{ m}
\end{aligned} \tag{2.34}
$$

where $\gamma = (f_{L1}/f_{L2})^2 = (77/60)^2$ is the ratio of the L1 to L2 carrier frequencies. For the dual-frequency user, this measure of P_{IF} is applied to eliminate the ionospheric propagation errors. The satellite and receiver TEC biases are eliminated when using this method (the satellite and receiver TEC biases are a constant that is added into the overall TEC value).

Dual frequency measurements are difficult to obtain because they either require a receiver capable of decoding the encrypted military signal, or they are obtained using semi-codeless methods, which significantly increases the measurement noise. These issues will be discussed in the coming chapters.

Most users of GPS utilize only the L1 signal upon which a ranging code is available for civilian access. This code will be discussed in more detail in Chapter 3. For the L1-only user, a model is broadcast by the satellites that is applied to partially correct for the ionosphere. As we shall soon see, however, this model is abstruse and generally inadequate.

An ionospheric model is broadcast as part of the navigation message from the satellites. The model approximates the ionosphere as a thin shell of ionization at a constant altitude of 350 km. The zenith delay at L1 is determined as (*IS-GPS-200D*, 2004):

$$T_{iono} = \left\{ \begin{array}{ll} F \times \left[5.0 \cdot 10^{-9} + \text{AMP}\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)\right] & , |x| < \pi/2 \\ F \times [5.0 \cdot 10^{-9}] & , |x| \geq \pi/2 \end{array} \right\} \text{ s} \quad (2.35)$$

where $\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)$ is simply the first three terms of the Taylor Series expansion of $\cos(x)$, and

$$\text{AMP} = \left\{ \begin{array}{ll} \sum_{n=0}^{3} \alpha_n \phi_m^n & , \text{AMP} \geq 0 \\ \text{if AMP} < 0 & , \text{AMP} = 0 \end{array} \right\} \text{ s} \quad (2.36)$$

and

$$x = \frac{2\pi\,(t - 50400)}{\text{PER}} \text{ rad} \quad (2.37)$$

where

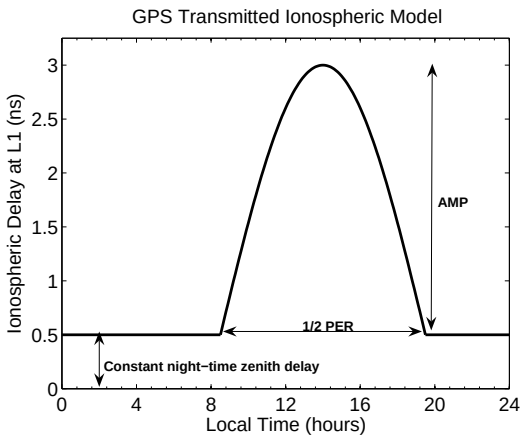

Figure 2.5: Example of the L1 GPS ionospheric model, illustrating the relationship between the parameters.

$$\text{PER} = \left\{ \begin{array}{ll} \sum_{n=0}^{3} \beta_n \phi_m^n & , \text{PER} \geq 72000 \\ \text{if PER} < 72000 & , \text{PER} = 72000 \end{array} \right\} \tag{2.38}$$

and finally, the obliquity factor, F, is:

$$F = 1.0 + 16.0\,[0.53 - E]^3 \tag{2.39}$$

α_n and β_n are the eight transmitted (four each) ionospheric parameters in the navigation message, and E is the elevation to the satellite in semi-circles (e.g. to convert from degrees to semi-circles, divide by $180°/\text{semi} - \text{circle}$). The AMP term determines the maximum amplitude of ionospheric delay, which occurs at a fixed local time (LT) of $t = 50400$ s (1400 LT). The PER term determines the width (period) of the cosine function. Finally, the obliquity factor, F accounts for the additional delay encountered by a low-elevation signal. Figure 2.5 illustrates the relationship between the parameter values, assuming an obliquity factor of $F = 1$.

The delay induced by the ionosphere depends on the point where the satellite signal traverses the ionosphere at the 350 km shell altitude on its way to the receiver. This point is known as the ionospheric puncture point. Moreover, the ionosphere is structured by the geomagnetic field, which, to first order, is tilted with respect to the Earth's rotational axis. The ionospheric model takes into account the ionospheric puncture point of the satellite signal as well as the geomagnetic field. The following equations develop these relationships:

$$\phi_m = \phi_i + 0.064 \cos\left(\lambda_i - 1.617\right) \text{ semi-circles} \tag{2.40}$$

$$\tag{2.41}$$

is the geomagnetic latitude of the earth projection of the ionospheric puncture point at an assumed height of 350 km, and

$$\lambda_i = \lambda_u + \frac{\psi \sin\left(A\right)}{\cos\left(\phi_i\right)} \text{ semi-circles} \tag{2.42}$$

is the geodetic longitude of the earth projection of the ionospheric puncture point at an assumed height of 350 km, A is the satellite azimuth in semi-circles measured clockwise from north, and λ_u is the receiver WGS-84 geodetic longitude in semi-circles. ϕ_i, the geodetic latitude of the earth projection of the ionospheric puncture point, is given as:

$$\phi_i = \left\{ \begin{array}{ll} \phi_u + \psi \cos\left(A\right) & , |\phi_i| \leq 0.416 \\ \text{if } |\phi_i| > 0.416 & , \phi_i = \text{sign}\left(\phi_i\right) 0.416 \end{array} \right\} \text{ semi-circles} \tag{2.43}$$

where ϕ_u is the receiver WGS-84 geodetic latitude in semi-circles, and sign refers to the signum function. ψ is the earth central-angle between the user position and the earth projection of the ionospheric puncture point and is given as:

$$\psi = \frac{0.0137}{E + 0.11} - 0.022 \text{ semi-circles} \tag{2.44}$$

Finally the local time, t, is:

$$t = \text{mod}\left(\lambda_i 4.32 \cdot 10^4 + t_{GPS}, 86400\right) \text{ s} \tag{2.45}$$

where t_{GPS} is the GPS time at the time of observation, and mod refers to the modulo function.

In Figure 2.6, the top panel illustrates the TEC of the ionosphere as estimated using the model described above. The bottom panel is a map of the true state of the ionosphere as computed using dual-frequency GPS receivers. Large discrepancies between the broadcast model and the true state of the ionosphere are apparent. In general, the model falls short for several reasons. The simplistic model cannot account for the complicated processes, dynamics, and structures often observed in the ionosphere even during geomagnetically periods or during solar minimum, as shown in Figure 2.6. Many features, such as the two large crests of ionization near the geomagnetic equator, are lacking and the densities in the model are underestimated. In other regions, the model overestimates the density of the ionosphere. The structuring of the ionosphere around the magnetic equator is also severely lacking. Finally, the obliquity factor (Equation 2.39) fails to adequately determine

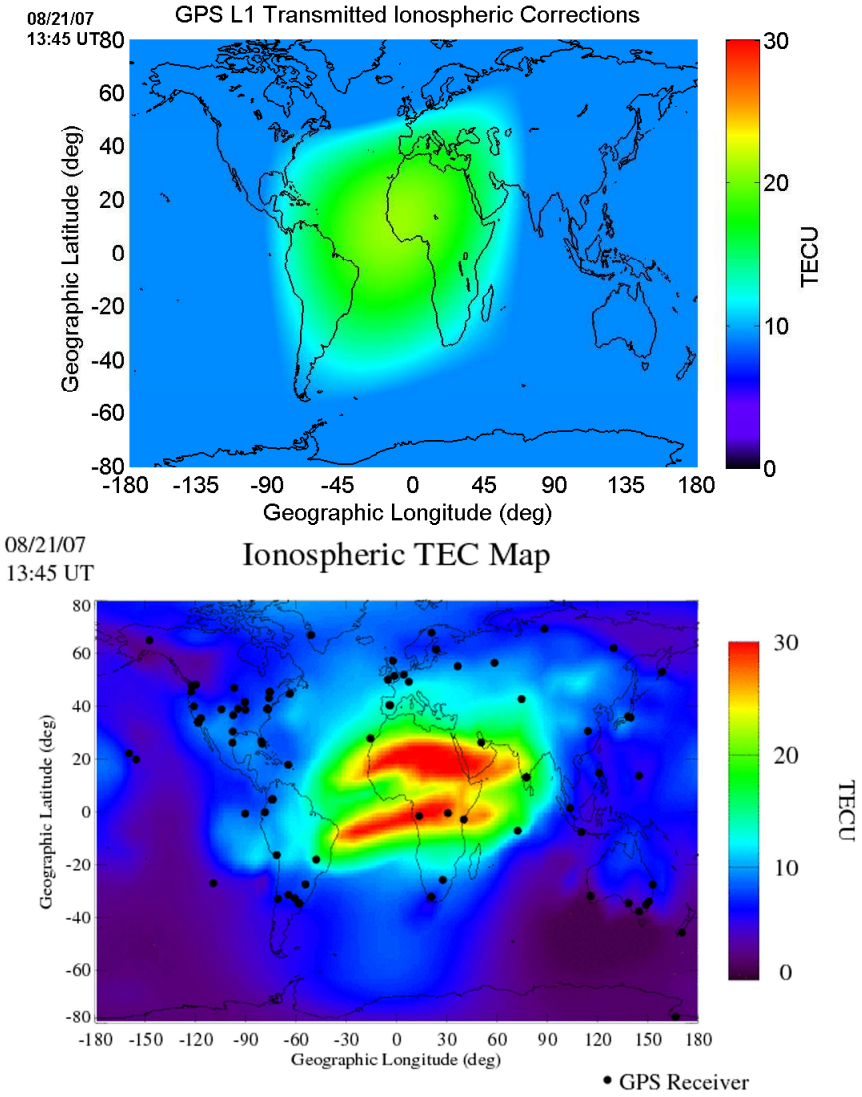

Figure 2.6: The top panel shows the vertical Total Electron Content of the ionosphere as estimated by the GPS transmitted model. The bottom panel is an estimate of the true state of the ionosphere created from world-wide dual-frequency observations of GPS satellites. The color scales are approximately equal. Even at solar minimum and during geomagnetically quiet periods, the discrepancies between the model and the true state of ionosphere are quite apparent. Bottom panel courtesy of NASA/JPL (http://iono.jpl.nasa.gov/latest_rti_global.html).

the TEC for low slant path angles, especially when the signal path traverses the equatorial ionization anomalies (the dense bands around the magnetic equator).

The model is deficient in other areas that are not well illustrated by the figure. For example, the peak ionization is not guaranteed to occur at 1400 LT, but instead can vary between 1100 and 1700 LT depending on the season, solar variability, etc. Moreover, the model falls far short during solar flare events or geomagnetic storms, when the ionosphere changes structure and density rapidly; these concepts are explored in more detail in Sections 2.3 and 2.5.

The inadequacies in the broadcast model result in residual ranging errors that can be positive or negative, depending on the state of the ionosphere and the model parameters. Although it is easy to criticize the model for its inadequacies, one must recognize that it has to be simple to accommodate the low 50 bit-per-second data transmission rate as well as the microprocessor capabilities available in the late 1970s when GPS was designed. Furthermore, until May 2000, the residual errors caused by the ionosphere were buried within the measurement noise due to purposeful degradation of the system by the Department of Defense in a feature known as Selective Availability. With the drastic increases in accuracy of the GPS system, the ionosphere has become the largest source of error for the single-frequency user.

2.3 Solar Flares

Most solar flares occur in active regions above sunspots and are characterized by drastic increases in x-ray, ultraviolet, and sometimes visible light. The energy they release is derived from magnetic reconnection, which occurs when twisted magnetic

fields above a sunspot rapidly break and reconnect (*Forbes and Acton*, 1996). During the reconnection process, protons, ions, and electrons are accelerated to relativistic velocities, and large, sudden increases of x-ray and ultraviolet radiation occur.

To measure the strength of a solar flare, earth-orbiting satellites are required because the ionosphere shields ground-based observatories from most of the x-ray and ultraviolet flux. Geostationary Operational Environmental Satellites (GOES) monitor not only Earth weather, but also the sun through an onboard magnetometer, x-ray sensor, and high-energy particle detectors[4]. GOES reports the soft (1.0 - 8.0 $\mathring{A}$) and hard (0.5 - 4.0 $\mathring{A}$) x-ray flux of the sun, which are classified based on their power density: A [10^{-8} - 10^{-7}) W/m^2, B [10^{-7} - 10^{-6}) W/m^2, C [10^{-6} - 10^{-5}) W/m^2. M [10^{-5} - 10^{-4}) W/m^2, and X [10^{-4} - 10^{-3}) W/m^2. For example, a flare of $2.5 \cdot 10^{-4}$ W/m^2 is given a classification of X2.5. The largest solar flare observed by GOES was measured at X28, but the device likely saturated, and a more likely value is X40 - X45 (*Thomson et al.*, 2005).

Solar flares have primarily one effect on space weather. The sudden increase in x-ray and ultraviolet radiation arrives at the Earth and immediately causes an increase in ionospheric plasma density. The increased ionization persists for the duration of the flare. The ionization can take up to two hours to completely return to pre-flare values, with lower regions of the ionosphere returning to pre-flare values more quickly due to higher recombination rates. Although the solar flare produces increased ionization across the sunlit hemisphere, the ionization is greatest at the subsolar point, where the incoming rays are perpendicular to the ionosphere and are thus most able to penetrate and produce the largest volume production rates.

[4]http://noaasis.noaa.gov/NOAASIS/ml/genlsatl.html

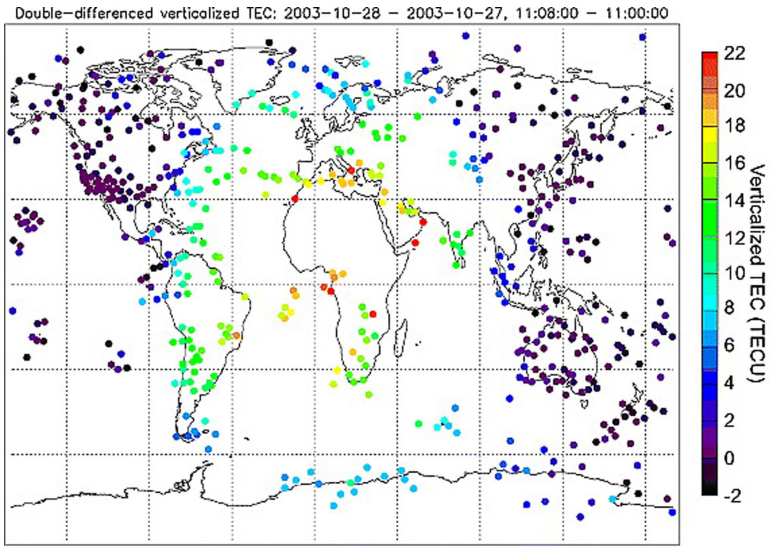

Figure 2.7: Vertical TECU enhancement over pre-flare ionosphere values for the X17 28 October 2003 solar flare. Data from approximately 100 dual-frequency ground-based GPS receivers are used to create the figure. The subsolar point is at the center of the figure, in Africa, and is concurrent with the location of the greatest TEC enhancement. Figure originally appeared in *Tsurutani et al.* (2005), © American Geophysical Union, reproduced with permission.

The x-rays and ultraviolet radiation create additional ionization at all levels of the ionosphere, but the radiation is capable of penetrating well into the *D*-region. *D*-region ionization increases substantially, causing high-frequency radio waves (∼3-30 MHz) to be drastically attenuated (*Mitra*, 1974).

Since solar flares suddenly increase the TEC between the satellite and receiver, the ranging error induced by the ionosphere is also increased. Figure 2.7 illustrates the increase in TEC for the 28 October 2003 X17 solar flare. At the subsolar point, this solar flare increased ionization by 30% (∼ 25 TECU) over pre-flare values (*Tsurutani et al.*, 2005). For dual-frequency users, the increase in TEC will not affect ranging precision because the receivers can correct the pseudoranges. Single-frequency users, however, are at a severe disadvantage. The broadcast model are not updated rapidly enough to account for the solar flare. Thus, ranging errors over the duration of the flare will be higher.

2.4 The Equatorial Ionosphere

In equatorial regions, GPS is primarily affected by the equatorial ionization anomaly (EIA) and a type of space weather known as convective ionospheric storms[5].

The EIA is a pair of increased ionization bands that reside approximately 15° off the geomagnetic equator on the sunlit side of the Earth. Day-time thermospheric winds generate eastward dynamo electric fields in the equatorial E-region that, aided by high conductivity, are transmitted along the geomagnetic field lines to the F-region. The plasma generated in the F-region is lifted to higher altitudes via an $\mathbf{E} \times \mathbf{B}$ plasma drift ($\mathbf{E}$ is eastward while $\mathbf{B}$, the geomagnetic field is horizontal and northward pointing over the equator). The plasma is lifted to high altitudes and eventually drifts down the geomagnetic field lines. The result is an accumulation of dense plasma roughly $\pm 15°$ from the geomagnetic equator. A small density trough typically exists between the anomaly peaks over the geomagnetic equator, and the density peaks may be asymmetric due to meridional (north-south) neutral winds.

The equatorial ionization anomaly is clearly visible as the parallel, high density plasma (red shades) swaths in Figures 2.4 and the bottom panel of 2.6. The main impact of the equatorial ionization anomaly is, once again, to single frequency users who are not completely able to correct for the ionosphere. Specifically, the peak density and large gradients associated with the anomalies are poorly modeled by the GPS broadcast model.

Convective ionospheric storms are caused by plasma instabilities that occur

[5]Convective ionospheric storms are commonly referred to as equatorial spread-F. Equatorial spread-F, however, is technically reserved for the radar and ionosonde echoes received from convective ionospheric storms.

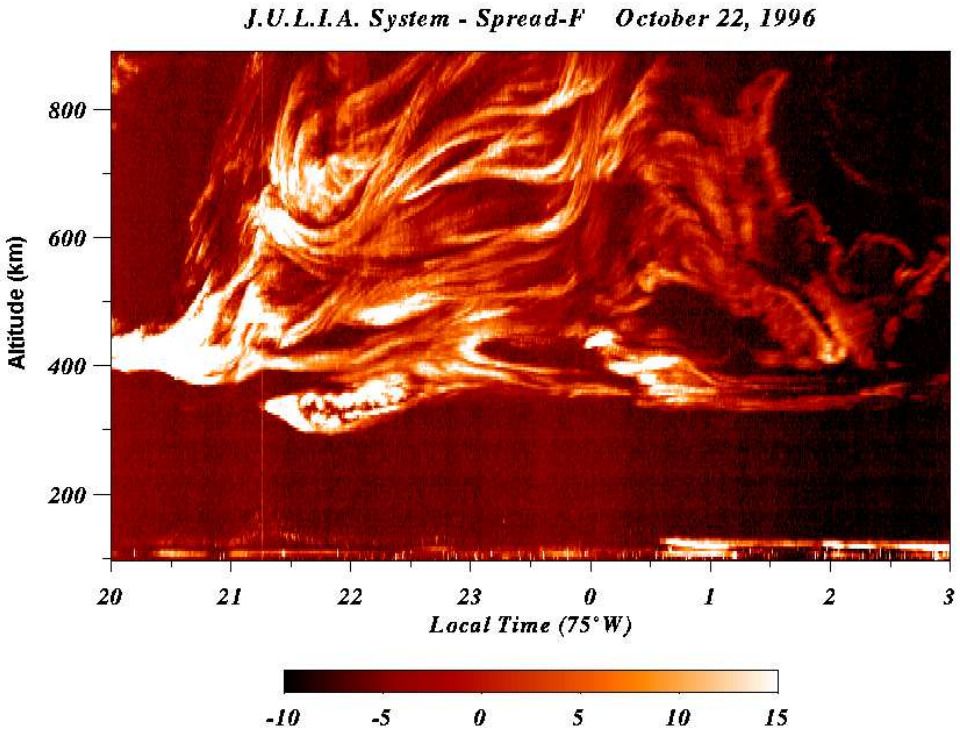

Figure 2.8: Coherent scatter return of a convective ionospheric storm as observed with the JULIA radar at Jicamarca Peru. The color scale is in dB. Figure originally appeared in *Kelley et al.* (2006), © American Geophysical Union, reproduced with permission.

only at night in the tropics. During the day, equatorial plasma is lifted to high altitudes by the E-region dynamo, and there is continuous replenishment of the plasma at lower altitudes by solar irradiation. As the ionosphere rotates into the dusk terminator, the E-region zonal winds blow more strongly and the E-region dynamo intensifies due to changes in conductivity, further lifting the F-region. After dusk, the E-region ionosphere rapidly recombines, while the raised F-region remains highly ionized. A steep vertical plasma density gradient occurs on the bottom of the F-region, creating the plasma equivalent of the Rayleigh-Taylor instability (the dense plasma is supported by the magnetic field above a layer of significantly lower plasma density).

A density perturbation in the form of gravity waves (*Nicolls and Kelley*, 2005)

or zonal shear (*Hysell and Kudeki*, 2004) perturbs the bottom of the F-region, triggering the convective ionospheric storm. The less dense plasma rises through the overlying, dense F-layer producing irregularities, sometimes known as plasma bubbles. Due to the high conductivity along the geomagnetic field lines in the F-region, the bubbles are elongated in the meridional (north-south) direction for many hundreds of kilometers, and the zonal scale sizes of the bubbles range from tens of meters to over a hundred kilometers. Convective ionospheric storms grow rather rapidly, with vertical bubble velocities on the order of 100 to 500 m/s, with some bubbles having velocities up to 5 km/s, with peak bubble altitudes sometimes as high as 1500 km (*Schunk and Nagy*, 2000a). After the growth period, the bubbles become 'fossilized', drift eastward with the surrounding plasma and slowly dissipate. Some bubbles persist until sunrise when sunlight finally ionizes the interior of the bubbles. A radar return of a convective ionospheric storm made using the JULIA coherent scatter radar at Jicamarca, Peru is shown in Figure 2.8.

Convective ionospheric storms do not occur every night in tropical regions. The peak altitude of the F-region and the plasma density gradient must be sufficient to trigger the storm. Moreover, their occurrence and severity is positively correlated to solar cycle variations, and in various regions of the world they occur seasonally. For example, in Hawaii, convective ionospheric storms occur more frequently during the equinoxes (*Makela et al.*, 2004). In South America, however, they are predominantly a northern winter phenomenon, and, in season, can occur as frequently as two out of every three nights.

One explanation for this behavior is the angle of the terminator with respect to the magnetic field lines: if one end of a magnetic flux tube is sunlit, the entire flux tube is shorted out into the E-region, effectively eliminating the formation of

bubbles. In Hawaii, the terminator and geomagnetic field lines tend to be aligned during the equinoxes. In South America, the declination of the geomagnetic field lines is large, and the lines are only aligned with the terminator during northern winter months (*Tsunoda*, 1985).

Convective ionospheric storms are also suppressed by meridional neutral winds, which tend to produce plasma density asymmetries that stabilize the Rayleigh-Taylor instability. The meridional neutral winds are stronger during the summer and winter months due to unequal solar irradiation of the northern and southern hemispheres. At the onset of a geomagnetic storm, convective ionospheric storms may be triggered by equatorward propagating gravity waves or by penetrating electric fields that can lift the ionosphere to great heights. After the onset of a geomagnetic storm, however, convective ionospheric storms tend to be suppressed due to decreased plasma densities.

The bubbles associated with convective ionospheric storms cause transionospheric radio waves to scintillate. As indicated by *Aarons* (1982), and as shown by *Rodrigues* (2003) for GPS, equatorial scintillations are most frequent and intense under the equatorial ionization anomalies and typically begin about 1 - 2 hours after sunset.

While passing through the ionosphere, the signal is diffracted by plasma density irregularities of the appropriate Fresnel scale size (approximately 250 m for GPS) embedded within the surrounding plasma of the ionosphere. The diffraction of the signal produces a scintillation pattern on the ground that varies both in amplitude and phase (*Yeh and Liu*, 1982). The amplitude scintillations are caused by phase-wise addition (diffractive process) of the GPS signal at the antenna and are characterized by rapid fluctuations both above (constructive interference) and

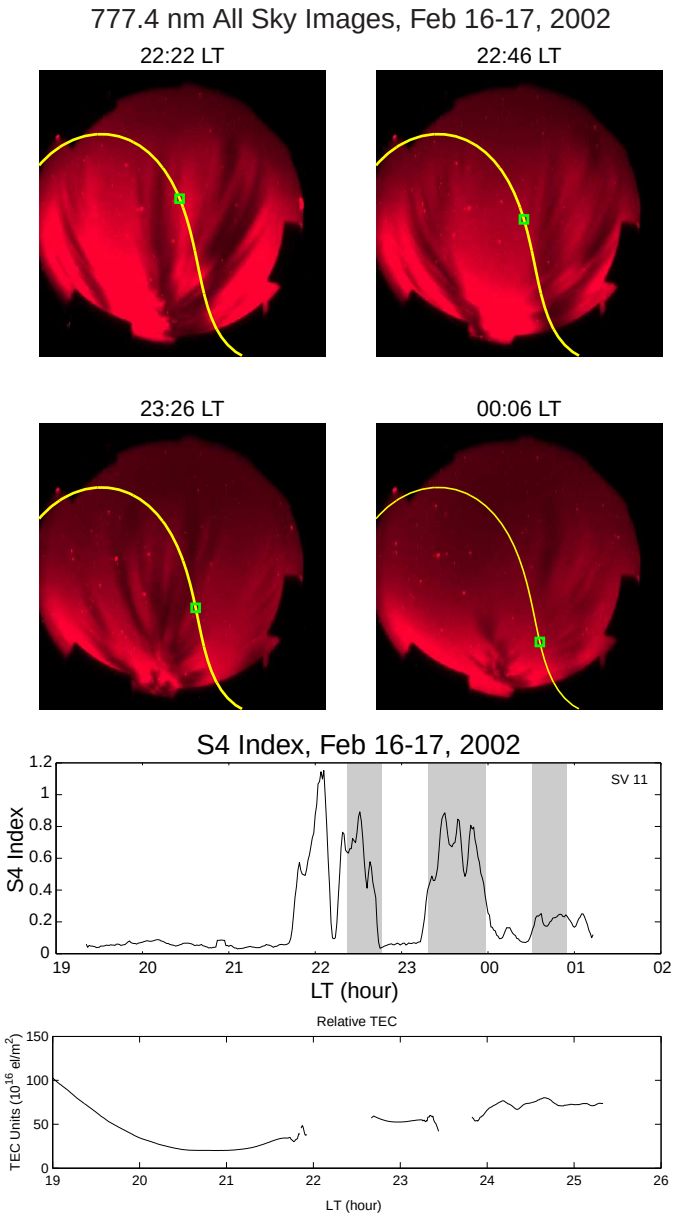

Figure 2.9: The top panel shows all-sky images of plasma bubbles over Hawaii, with the trajectory of a GPS satellite (yellow). The square on each track denotes the location of the satellite when the image was taken. As the satellite signal traverses the bubbles, the signal begins to scintillate, resulting in elevated S_4 values (middle panel). The shaded regions correspond to times when the look direction to the satellite traversed the depleted regions in the images. Dual-frequency TEC measurements from a semi-codeless NovAtel receiever (bottom panel) are difficult to obtain because semi-codeless GPS receivers are prone to loss of lock during scintillation. Figure originally appeared in *Kelley et al.* (2002), © American Geophysical Union, reproduced with permission.

below (destructive interference) the received carrier power from the satellite. Phase scintillations are also produced by the interference patterns as well as changes in TEC (refractive process) as the signal traverses through the bubbles of depleted electron density. Figure 2.9 shows a convective ionospheric storm bubble visible over Hawaii. The bubbles induced TEC variations and GPS amplitude scintillation, which is indicated by the high S_4 values.

The S_4 index is a common metric to measure the severity of the amplitude scintillation and is defined as (*Yeh and Liu*, 1982):

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}} \tag{2.46}$$

where I is the intensity of the signal. The term under the radical is merely the standard deviation of the signal amplitude divided by the mean signal amplitude. The S_4 index has a maximum value of 1.2. Typically, for GPS, the intensity, or carrier-to-noise ratio, is measured at 50 samples per second, and the S_4 index is averaged over a one-minute period. For GPS, S_4 indices from about 0.15 - 0.4 are considered weak scintillations, from 0.4 - 0.7 are considered moderate, and from 0.7 - 1.2 are considered severe.

The time-scale of the fades depends on the Fresnel scale size as well as the bubble drift velocity (typically 100 - 200 m/s) and is complicated by the fact that the GPS signal puncture point is also moving across the ionosphere. Typical fading time-scales for GPS are from 0.5 - 2 seconds (*Humphreys et al.*, 2007), but can be much longer ($\sim$ 10 s) when the GPS signal velocity at the ionospheric signal puncture point matches the ionospheric zonal velocity (*Kintner et al.*, 2004).

The amplitude fades can be quite large, in some cases in excess of 30 dB

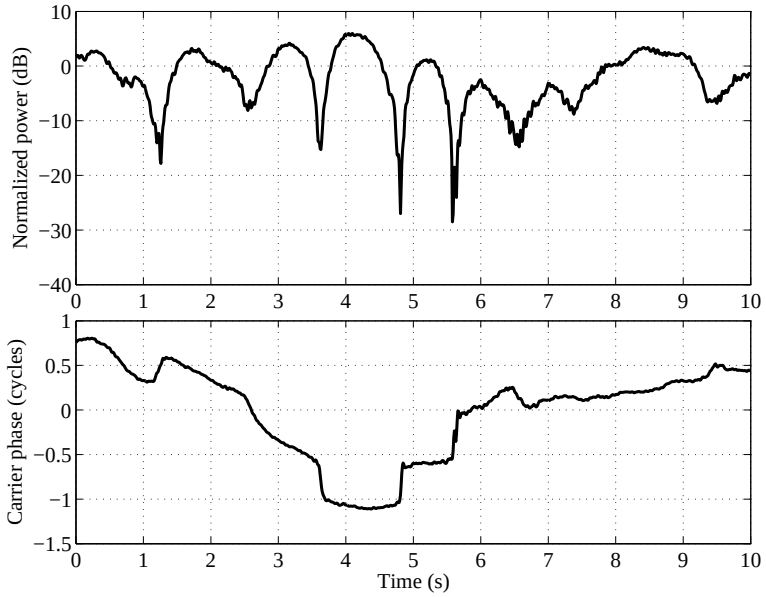

Figure 2.10: Figure illustrating 'canonical fades' associated severe equatorial amplitude and phase scintillation. The canonical fades are the deep, wedge shaped amplitude fades (upper panel) at approximately 3.7, 4.8 and 5.6 s that are accompanied by approximately 1/2 cycle abrupt phase changes (lower panel). The data was acquired in Cachoeira Paulista, Brazil (22.7°, 45.0°W) with $S_4 \approx 0.9$. Figure originally appeared in *Humphreys et al.* (2007). © IEEE, reproduced with permission.

(*Humphreys et al.*, 2005). The largest fades ('canonical fades') are associated with abrupt, approximately half-cycle phase changes and are a feature of strong equatorial scintillation (*Humphreys et al.*, 2007). Examples of such fades are shown in Figure 2.10. Scintillation-inducing irregularities may occupy the entire sky and broad geographic areas. Therefore, it is not uncommon for a receiver to view many satellites that are simultaneously scintillating, as shown in Figure 2.11.

Scintillations are more intense at lower frequencies, so the GPS L2 frequency is more susceptible to scintillation than the L1 frequency (*Yeh and Liu*, 1982). For weak scintillations, the amplitude fades are time-correlated, but for strong scintillations, this scenario is not true (*Humphreys et al.*, 2007). Figure 2.12 illustrates weak-moderate scintillations ($0.2 \leq S_4 \leq 0.5$) at L1 and L2 using a coded civilian

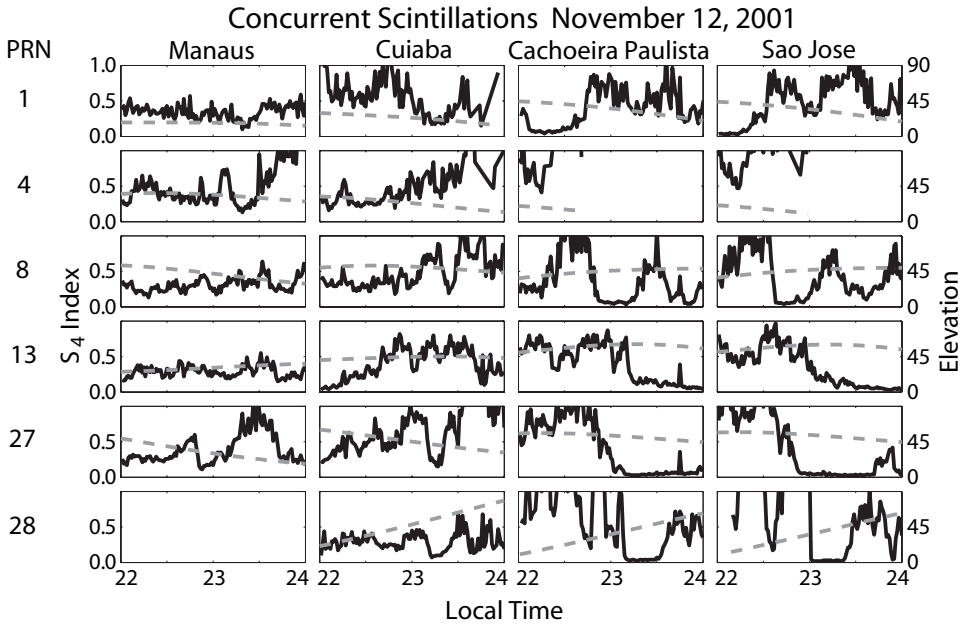

Figure 2.11: Figure illustrating an active evening of scintillation in Brazil during solar maximum. The majority of satellites tracked by each receiver were scintillating simultaneously. The elevation of each satellite is plotted as a dashed line. The location of each receiver is: Manaus (3°S, 60.0°W), Cuiaba, (15.6°S, 56.1°W), Cachoeira Paulista (22.7°, 45.0°W), and Sao Jose dos Campos (23.2°S, 45.9°W). Figure originally appeared in *Kintner et al.* (2007). © 2007 American Geophysical Union, reproduced with permission.

receiver in Natal, Brazil (5.8°S, 35.2°W). For these scintillations, the fades are time-correlated, however, in general, the L2 amplitude fades are deeper than the L1 fades, as expected.

All users (civilian, military, single or dual-frequency systems) are vulnerable to scintillation. The deep fades and abrupt phase changes associated with severe scintillations are particularly difficult for the receiver tracking loops to maintain carrier lock on the satellite signal. Dropping satellites results in increased dilution of precision, and in extreme cases may lead to loss of navigation solution. Civilian semi-codeless dual frequency receivers are particularly prone to scintillations on the L2 frequency because of the semi-codeless tracking methods that are utilized. Losing carrier lock is particulary detrimental to carrier-phase differential

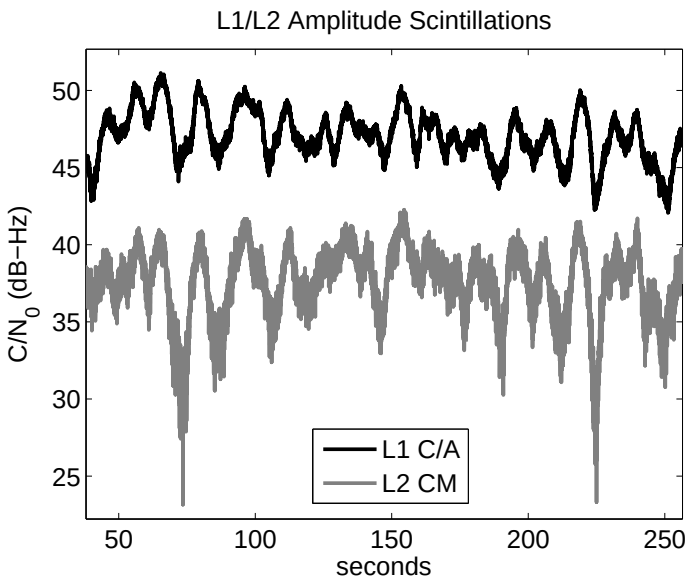

Figure 2.12: Weak-moderate equatorial scintillations recorded in Natal, Brazil (5.8°S, 35.2°W) in January 2007 on PRN 12 using the civilian L1 C/A and L2 CM signals. The scintillations were recorded using a dual-frequency bit-grabber and processed in Matlab.

GPS applications that require precise, uninterrupted carrier phase information for sub-meter positioning.

Testing receivers under scintillating environments is difficult because real scintillations are difficult to predict and repeatable scenarios are not available. Fortunately, much work in developing a realistic scintillation test bed for use in conjunction with GPS simulators has been recently developed by *Humphreys et al.* (2007). Few GPS engineers have tackled the difficult subject of improving receiver performance in scintillating environments. It is beyond the scope of this chapter to discuss various mitigation techniques, but the author suggests *Van Dierendonck* (1998) and *Humphreys et al.* (2005) as a starting point. *Kintner et al.* (2007) provide a complete treatise on GPS and scintillations.

2.5 Geomagnetic Storms and the Mid-latitude Ionosphere

The mid-latitude ionosphere is typically uneventful except during geomagnetic storms. Therefore, in this section, we quickly begin with geomagnetic storm-time effects on the ionosphere, and move on to uniquely mid-latitude space weather phenomenon that affect the Global Positioning System.

Geomagnetic storms are triggered by shocks in the solar wind that compress the magnetosphere. A major mechanism of energy transfer from the solar wind to the Earth's magnetosphere is through magnetic reconnection (*Dungey*, 1961). When the interplanetary magnetic field carried by the solar wind has opposite polarity (southward) to that of the Earth's field (northward), the Earth's magnetic field facing into the sun is stripped away and reconnected deep in the nightside magnetotail via magnetic reconnection (*Tsurutani et al.*, 2003). In the magnetotail, the reconnection processes accelerate particles earthward that precipitate into the polar atmosphere to create the aurora and inject the high-latitude ionosphere with enormous amounts of energy. Furthermore, highly energetic protons and electrons form a dense ring current in the Van Allen radiation belts with a peak density around 3 - 4 R_E[6]. The ring current causes a depression of the magnetic field intensity at the equator, and the departure of the magnetic field intensity from quiet conditions is a measured index known as D_{st}, which is an excellent indicator of the strength of the storm (*Dessler and Parker* (1959), *Sckopke* (1966)). A large negative decrease of the D_{st} index marks the commencement of a geomagnetic

[6]The GPS satellites reside at $\sim 4R_E$ orbits, placing them in the Van Allen radiation belts. The highly-accelerated protons and electrons primarily affect the GPS satellites by decreasing the life-span and efficiency of the solar panels. To date, there are no (public) reports of a GPS satellite having been directly disabled by a magnetic storm event.

storm.

Although storm periods are categorized by names (e.g. storm commencement, initial phase, main phase, and recovery phase), the response of the ionosphere can be summarized as a positive phase and negative phase. An excellent review of storm-time processes and the associated TEC variations is found in *Mendillo* (2006). The response of the positive phase is associated with enhanced global TEC and equatorward motion of polar ionospheric features, and poleward motion of low latitude features. In response to the high-energy input at polar latitudes, the auroral E-region electron densities increase and the ion and electron temperatures increase (*Schunk and Nagy*, 2000a). The auroral oval is driven equatorward[7]. The neutral composition of the ionosphere also changes, increasing the neutral winds, and triggering equatorward propagating gravity waves (*Valladares et al.*, 2007).

At low latitudes, the response of the ionosphere during the positive phase is an enhancement of the equatorial ionization anomaly, reducing the TEC directly over the equator, and driving the anomalies poleward and to extraordinary heights (in excess of 840 km in one case (*Greenspan et al.*, 1991)). The short-duration enhancement is triggered by the prompt appearance of electric fields within the magnetosphere that are projected onto the ionosphere. The extraordinary height of the ionosphere can lead to spectacular convective ionospheric storms that are visible as far north as Arecibo, Puerto Rico (*Makela et al.*, 2001), and Hawaii (*Kelley et al.*, 2002).

According to *Mendillo* (2006), the positive phase of the storm lasts longer and is more prominent at low latitudes. The positive phase of the storm typically has a maximum amplitude at 1800 LT at mid-latitudes. At polar latitudes, the

[7]In a famous superstorm of 1859 the aurora were visible down to latitudes within $\pm 23^{\circ}$ geomagnetic of the equator (*Tsurutani et al.*, 2003))

peak amplitude of the storm is weaker and occurs earlier. At low latitudes, the peak amplitude of the storm is also weaker and tends to occur later. Finally, positive storms are more pronounced in the winter hemisphere, and negative storms are more pronounced in the summer hemisphere. The positive phase of a storm typically lasts about a day, and is followed by the negative phase that has durations on the order of a few days.

The negative phase is typically associated with a global depression of overall TEC values. The depression of TEC is primarily caused by meridional neutral winds that change the neutral composition of the F-region of the ionosphere (*Strickland et al.*, 2001). The overall effect is to decrease the electron density, and a few days elapse before the ionosphere recovers pre-storm levels. In general, during the negative phase, convective ionospheric storms are suppressed because of the decreased electron density.

Figure 2.13 is from *Mendillo* (2006) and illustrates the percentage changes in TEC caused by the TEC records for 70 storms from Narssarssuaq (61.2°N, 45.4°W), 67 storms from Goose Bay (53.3°N, 60.4°W), 109 storms from Sagamore Hill (42.3°N, 70.9°W), and 70 storms from Cape Kennedy (28.5°N, 80.6°W). Notice the positive phase of the storms lasts about one day, followed by the negative phase of the storms that lasts roughly three days.

The large departures in TEC are, once again, not problematic for dual-frequency GPS users. However, the broadcast model for the single-frequency user cannot account for the sudden changes that accompany the onset of a geomagnetic storm. During these times, ranging errors will likely increase as the broadcast model will underestimate the TEC. Likewise, the intensity of the negative phase of the storm is also difficult to model and further ranging errors will ensue during

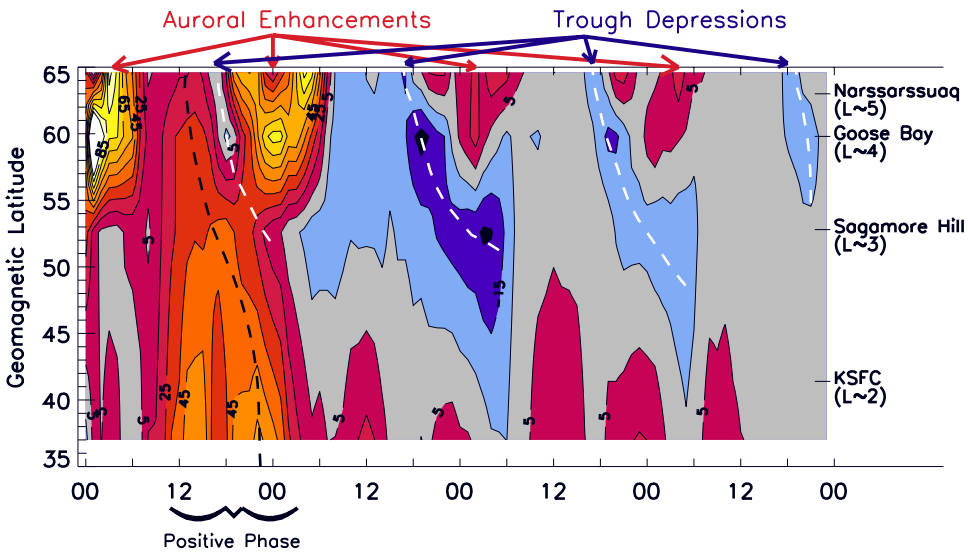

Figure 2.13: Figure illustrating the storm patterns for TEC percentage changes versus geomagnetic latitude (geographic latitude of the stations is in the text). Figure originally appeared in *Mendillo* (2006), © American Geophysical Union.

this period.

The model broadcast model is not updated fast enough nor detailed enough to account for the dynamic processes that are occurring in the ionosphere. Ranging errors are extreme during geomagnetic storms, and can be exacerbated by the GPS ionospheric model. Furthermore, although our understanding of the storm-time processes is advanced, our ability to model them is not very good, preventing the determination of more detailed corrections.

All users of GPS in the equatorial ionosphere will be prone to scintillations, and during the positive phase of the storm these scintillations extend poleward, affecting midlatitude users. The equatorward extension of the auroral oval is not detrimental since, as will be discussed in Section 2.6, the polar ionosphere generally does not pose many challenges to GPS receivers. In one case, *Sojka et al.* (2004)

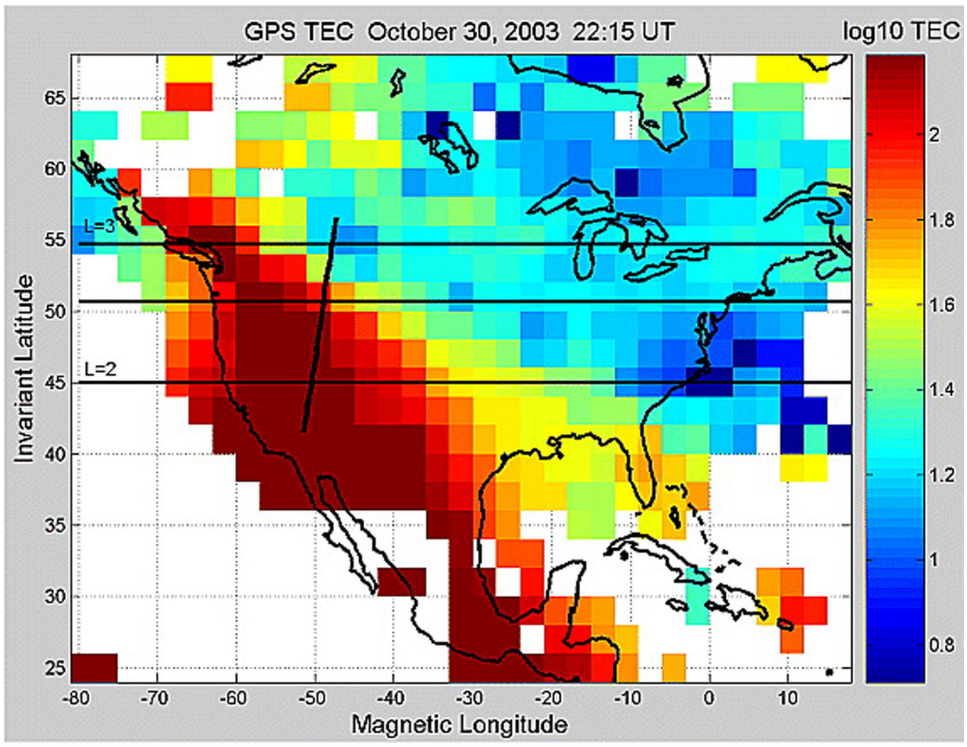

Figure 2.14: Figure illustrating an SED plume associated with the Halloween events of 2003. The plume exceeded 250 TECU over the northeast United States. Figure originally appeared in *Foster and Rideout* (2005), © American Geophysical Union.

observed scintillations of $S_4 = 0.18$ caused by auroral arcs with a receiver located in Utah.

The space weather effects on the mid-latitude regions that are considered here are not concerned with the equatorward movement of the aurora or the poleward movement of the equatorial ionization anomaly and convective ionospheric storms. Scintillations at mid-latitude that are not tied to these events appear to be associated with storm-time ionospheric phenomena (*Basu et al.*, 2005b). These events are commonly given names such as storm enhanced density (SED) or subauroral polarization streams (SAPS). A few theories have been proposed as to their origin and creation (e.g. *Mishin et al.* (2003b), *Mishin et al.* (2003a), and *Keskinen et al.* (2004)). Both cause scintillations and large TEC gradients that have been pre-

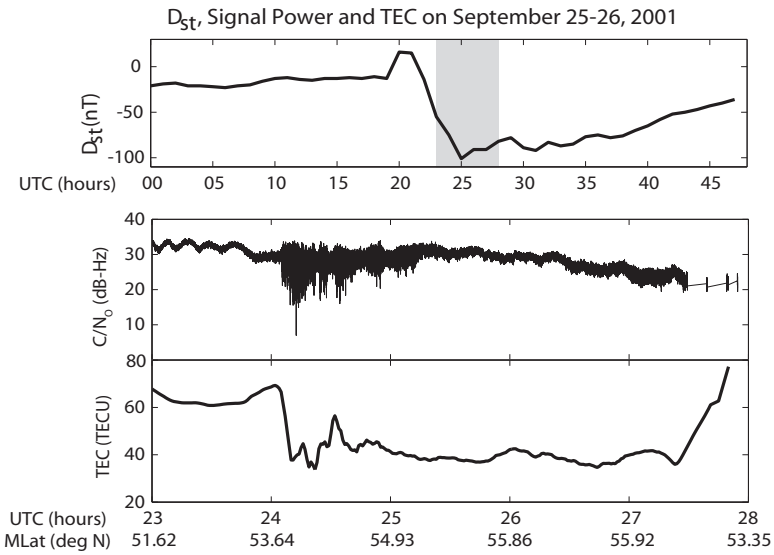

Figure 2.15: Top panel: D_{st} index for 25-26 September 2001. The shaded portion in the top panel refers to the times indicated in the middle and bottom panel. Middle panel: Amplitude scintillations observed at Ithaca, NY (42.4°N, 76.5°W). Bottom panel: Large TEC gradients that induced the scintillation. Figure originally appeared in *Ledvina et al.* (2002), © American Geophysical Union.

dominantly observed over the North American sector (e.g. *Ledvina et al.* (2002), *Foster et al.* (2007), and *Foster and Rideout* (2005)) and perhaps elsewhere (*Coster et al.*, 2005). SEDs are plumes of extremely dense plasma (in excess of 200 TECU) that appear to convect poleward from equatorial regions and in some cases convect over the polar cap (*Mitchell et al.*, 2004). An example of an SED plume is shown Figure 2.14. Figure 2.15 illustrates the severe ($S_4 = 0.8$) scintillations that occurred over Ithaca, NY (42.4°N, 76.5°W) coincident with a geomagnetic storm and strong TEC gradients (*Ledvina et al.*, 2002).

Naturally, these plumes, which are not predictable, add significant errors to single-frequency users. Single-frequency and dual-frequency users are also prone to the scintillations that occur on the steep TEC gradients at the edges of the plumes. The steep ionospheric gradients also pose challenges for differential GPS systems such as the Federal Aviation Administration's (FAA) Wide Area Augmentation

System (WAAS). The spacing of the reference receivers cannot adequately model the steep TEC gradients associated with the plumes. During large geomagnetic storms, the steep gradients impact WAAS availability (*Doherty et al.*, 2003).

2.6 Polar Ionosphere

In general, the polar ionosphere does not possess large enough densities to induce strong amplitude scintillations, except in rare cases. Irregularities do exist in polar regions and are responsible for creating phase scintillations. Although these phase scintillations have been strong enough to cause cycle slips, GPS does not appear to be nearly as prone to loss of lock as is the case for equatorial and mid-latitude scintillations.

Intense diffractive scintillations have been observed by *Basu et al.* (1998) that are caused by F-region irregularities in polar cap patches (*Coker et al.*, 2004). These irregularities do not have sufficient density to cause severe amplitude scintillations at GPS frequencies (*Kintner et al.*, 2007). As reported by *Aarons* (1997), however, phase scintillations are quite common in the polar ionosphere. *Basu et al.* (1998) reported scintillations by observing the time rate of change of TEC and observed changes of roughly 2 TECU/min.

Auroral arcs have been observed to cause strong scintillation on GPS signals (*Smith et al.*, 2007). Cycle slips and phase scintillations associated with GPS signals traversing particularly bright auroral arcs have been documented in the literature (e.g. *Kintner et al.* (2002) and *Kintner et al.* (2007)) and are more common than amplitude scintillations. An example of TEC fluctuations and a cycle slip induced by auroral arcs are shown in Figure 2.16. *Kintner et al.* (2007)

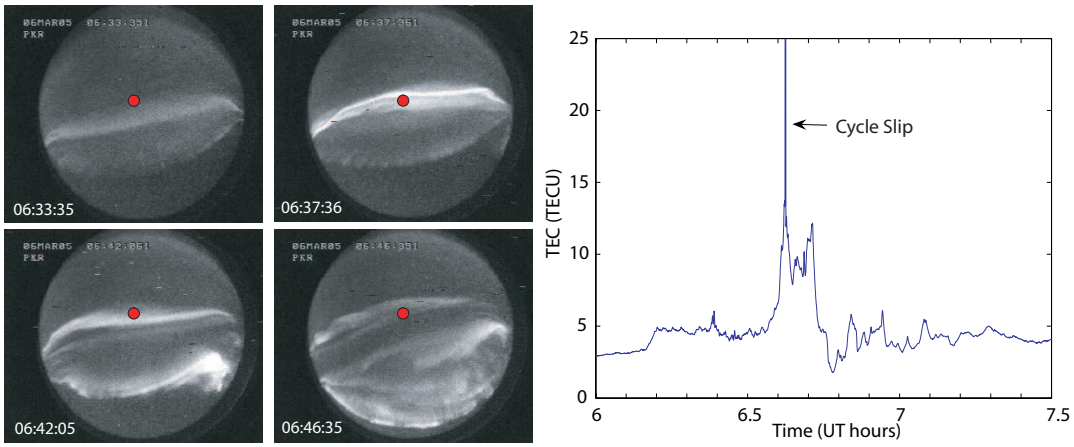

Figure 2.16: Left panel: All-sky camera pictures of an auroral breakup on 6 March 2005. The red dot represents the elevation and azimuth of a GPS satellite. South is at the top, and east is to the right. Right panel: TEC measured from the GPS satellite indicated in the left panel. The period of enhanced TEC measurements corresponds to the period of time spanned by the images in the left panel. Figure originally appeared in *Kintner et al.* (2007), © American Geophysical Union.

suspects that the cycle slip was caused by a rapid change in TEC on a step density gradient produced by the arc itself. A receiver capable of recording amplitude scintillations was not available, so the theory has yet to be verified.

The ionospheric densities over the polar cap are typically not high enough to induce the amplitude scintillations typically seen with gradients and convective ionospheric storms at middle and low latitudes. One exception appears to have occurred on the evening of 30 October 2003, when a large plume of plasma convected from the American sector, over the polar cap region, and into northern Europe and was documented by *Mitchell et al.* (2004). A GPS receiver located in Svalbard, Norway (78.9°, 11.9°E) recorded amplitude scintillation levels up to $S_4 \simeq 0.25$.

2.7 Solar Radio Bursts

Solar radio bursts are intense radio emissions from the sun that are usually associated with solar flares (*Kundu*, 1965; *Bastian et al.*, 1998). Unlike other space weather affecting GPS, they are not related to processes occurring in the ionosphere. Solar radio bursts are sudden, drastic increases in radio emissions from the sun.

Typically, incoherent gyrosynchrotron emissions dominate as the cause of these radio emissions at centimeter and millimeter wavelengths, whereas coherent plasma radiation is often the cause at meter and decimeter wavelengths. Solar radio bursts can also be created via a natural Electron-Cyclotron Maser emission process (*Treumann*, 2006). Similar to flares, the rate of occurrence of solar radio bursts follows the 11-year cycle of solar activity, being both more numerous and generally stronger near solar maximum. Solar radio burst strength (flux density) is measured in solar flux units (SFU; 1 SFU = 10^{-22} W m^{-2} Hz^{-1}), and is reported at various frequencies by several observatories around the world. As shown by *Nita et al.* (2002), solar radio bursts typically exhibit a power-law distribution of time between bursts versus peak intensity.

Plasma radiation occurs when electron beam energy is nonlinearly converted to plasma waves at harmonics of the local plasma frequency, which are then converted to ordinary-mode electromagnetic waves. Plasma radiation is primarily the source of solar radio bursts in the frequencies that will affect GPS. In a solar flare, large numbers of electrons are accelerated from the flare site and interact via electron beam instability with the surrounding local plasma and magnetic fields. Through a mode conversion process, the energy is converted into ordinary-mode electromagnetic radiation at a frequency equal to the local plasma frequency. This

electromagnetic radiation is the solar radio burst. Since the local magnetic field has no preferred direction, the electromagnetic wave arriving from the solar radio burst can be left-hand circularly polarized, right-hand circularly polarized, or a mixture of both with respect to its wave vector at Earth. Typically, these radio bursts occur in the frequency range from a few tens of megahertz up to 3 GHz (*Bastian et al.*, 1998).

Solar radio bursts can have durations from tens of seconds to a few hours. Typically, solar radio bursts that are concurrent with a solar flare are impulsive, while those that occur in the decay phase are smoother and of longer duration. Over the duration of the burst, the peak frequency can climb, fall, or both. Furthermore, the intensity and polarization of the burst can vary. Finally, the bandwidth of the solar radio burst at any point in time can vary substantially, and has been measured as narrow at 17 MHz (*Csillaghy and Benz*, 1993).

When a solar radio burst of the appropriate frequency arrives at the receiver, the signal appears as in-band interference. Since the Global Positioning System was designed as a right-hand circularly polarized (RHCP) system, only RHCP solar radio bursts affect the receiver. The in-band interference decreases the received carrier-to-noise ratio of every tracked satellite signal by increasing the noise floor of the receiver. The decreased carrier-to-noise ratio leads to loss of precision, or, in drastic cases may cause complete loss of lock on satellites, potentially leading to loss of navigation solution. The intensity of the fade, coupled with the gain of the GPS receiving antenna and the noise figure of the receiver determine the impact a given solar radio burst will have on carrier-to-noise ratio.

In the upper panel of Figure 2.17, the solar radio burst power spectrum from 1 to 2 GHz over the period of 18 UT to 20 UT on 6 December 2006 is shown. The

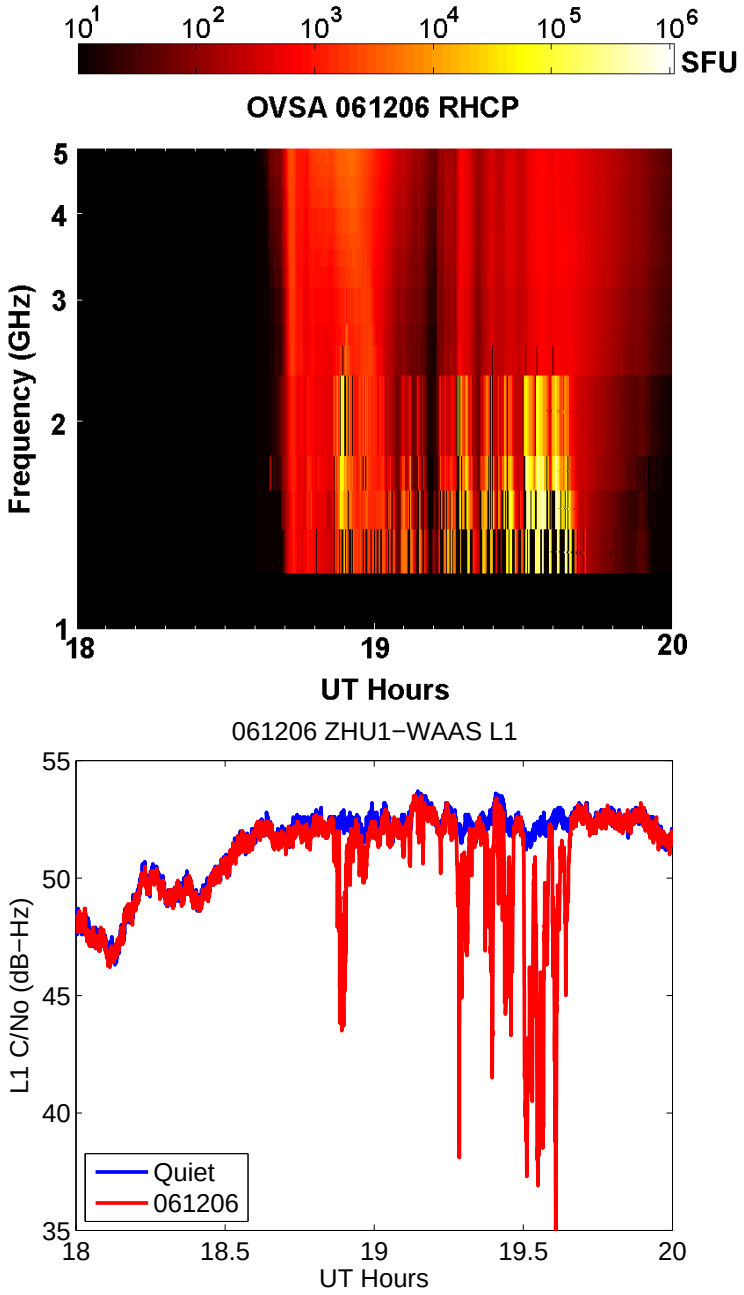

Figure 2.17: Top panel: Solar radio burst flux on 6 December 2006 as measured at the Owens Valley Solar Array. Bottom panel: Response of a GPS receiver to the solar radio burst. The red line corresponds to C/N_0 on 6 December 2006 and the blue line corresponds to the previous sidereal day. Figure originally appeared in *Cerruti et al.* (2007), © 2007 American Geophysical Union.

lower panel shows in red the values of C/N_0 as measured on a GPS receiver located in Houston, TX for 6 December 2006. The blue trace is the C/N_0 values for the same receiver for the previous sidereal day. The variation of both the red and blue values before 18:45 UT, when they vary identically, is the result of multipath interference originating near the receiving antenna. After about 18:55 UT the 6 December data show rapid decreases in C/N_0, corresponding to increases in solar radio burst power near 1.6 GHz. The 6 December 2006 solar radio burst was a record setting event in the GPS frequencies, with a flux density in excess of 500,000 SFU at 1.6 GHz. Prior to the events of December 2006, the record radio burst near the GPS frequencies, according to reports collected by the National Oceanic and Atmospheric Administration (NOAA), was 165,000 SFU at 1415 MHz for a SRB in April 1973. Second place was 88,000 SFU at 1415 MHz in February 1979. At this point, we shall leave the subject of solar radio bursts and their effects on GPS receivers as they are the scope of the rest of this thesis.

2.8 Conclusion

In the preceding pages several space weather effects on the Global Positioning System have been detailed. Most of the space weather effects on GPS are caused by the ionosphere, which is a dispersive medium. A dispersive medium induces a delay on a signal propagating through it that is proportional to the signal's frequency.

For single-frequency users of GPS, the ionosphere often presents the largest source of error, even with the mediocre model transmitted by the GPS satellites. The errors are largest for users in the tropics because of the dense ionosphere and

steep gradients associated with the equatorial ionization anomaly. The models cannot account for the abrupt, large-scale changes of the ionosphere during geomagnetic storms. Solar flares cause increases in TEC of up to 30% over pre-flare values and induce further positioning errors for single-frequency users (*Tsurutani et al.*, 2005).

In general, dual-frequency GPS users are generally not vulnerable to changes in TEC resulting from the equatorial ionization anomaly, solar flares, or geomagnetic storms because of their ability to correct for the dispersive propagation of the radio wave through the ionosphere. Dual-frequency receivers, however, are expensive and generally not as robust because they must track the encrypted military signal with civilian semi-codeless receivers. Users of the encrypted military code have full access to the signal and are not nearly as prone to loss of lock as the semi-codeless dual frequency users.

Of particular concern for GPS users of all types of receivers at low latitudes are convective ionospheric storms, which cause the satellite signal to scintillate both in phase and amplitude. These scintillations are difficult for the receiver to track through, leading to loss of lock on the satellite signal, which causes increased dilution of precision and positioning errors. If enough satellites are lost, then total loss of navigation solution occurs.

Solar radio bursts are a space weather phenomenon that affect all users of GPS receivers. Solar radio bursts are sudden increases of radio energy from the sun that interfere with the receiver's ability to track the GPS satellite signal. Solar radio bursts were previously not thought to affect GPS receivers (*Klobuchar et al.*, 1999), but recent events have challenged these notions. The impact of solar radio bursts on GPS receivers is the focus of the rest of this dissertation.

Chapter 3

Global Positioning System: The System, Signals, and Receivers

navigation (nav'ə gā'sh͡ən), *n.* **1.** the act or process of navigating. **2.** the art or science of plotting, ascertaining, or directing the course of a ship or aircraft. *Webster's Encyclopedic Unabridged Dictionary of the English Language*

3.1 Introduction

The purpose of the Global Positioning System (GPS) is to provide precise positioning, velocity and timing (PVT) information to its users. Space weather affects the receiver's ability to provide accurate PVT information. The goal of this chapter is to motivate an understanding of GPS in an operational sense. The foundation will be necessary for later understanding how solar radio bursts affect GPS.

The chapter is organized as follows. First, an overview of current GPS operations and systems requirements, including error budgets, will be given. Next, the current GPS signal structure as transmitted will be detailed, including both civilian and military codes. A discussion on received signal power will motivate an analysis of the signal operating environment. Next, a brief introduction to receiver technology will be given. Down-conversion, digitization, and link margins of the signal will be discussed. The chapter will conclude with a discussion of code-stripping and baseband mixing that will ultimately lead to navigation solution algorithms and a discussion on dilution of precision.

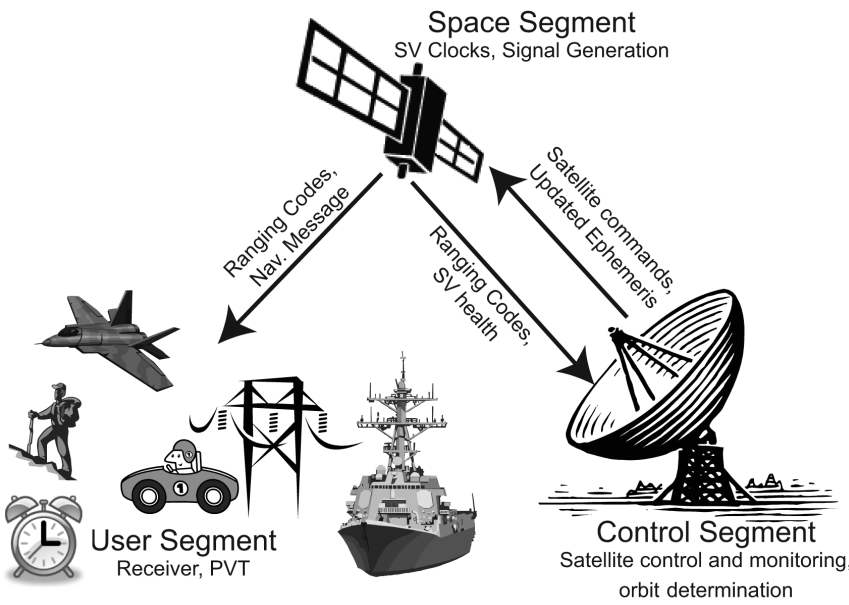

Figure 3.1: A cartoon illustrating the relationship between the three segments of the Global Positioning System.

3.2 Global Positioning System Operations

The Global Positioning System is divided into three segments: the space segment, the control segment, and the user segment and are shown in Figure 3.1. Together these three segments constitute GPS Operations. The space segment consists of at least 24 satellites that continuously transmit a ranging signal and data bits that include the satellite ephemeris and clock parameters. The control segment is responsible for tracking each satellite, monitoring its health, and uploading the satellite ephemeris and clock corrections. Finally, the user segment is any user that requires GPS position, velocity, and timing information. In the sections below, each segment is further detailed.

3.2.1 Space Segment

For a fully operational system, Congress mandates that at least 24 operational satellites be available 95% of the time over any day, with at least 21 transmitting satellites available 98% of the time in a year. As of May 2007 the constellation has 31 operational satellites. The satellites are launched into six orbital planes with a nearly circular orbital radius of 26,561.75 km. The ascending nodes of the six orbital planes are equally spaced around the earth, with at least four satellites unevenly spaced in each plane. The inclination of the orbital plane is 55 degrees, and a satellite completes an orbit in 12 sidereal hours. Thus, every 24 sidereal hours, the satellites occupy the same location in the sky as viewed from an observer on earth.

With these orbital parameters each satellite almost has a hemispherical view of the earth. The poles are covered simply by having a moderate inclination of 55 degrees, and complete earth coverage is guaranteed with a small number of satellites. Finally, at these altitudes atmospheric drag is small, allowing for precise orbital determination. Unfortunately, the satellites lie within the radiation belts and so spacecraft radiation and charging is significant. Ultimately, the choice of altitude provides the best compromise between lifetime, launch weight and cost, and global coverage.

As a result of the satellite placement, most earth-bound receivers typically see at least 6-8 satellites at any given time, and many locations on earth frequently see 10 or more satellites for portions of the day. The orbital inclination and satellite placement within the orbits provides good geometric diversity. Increasing the number of satellites in the constellation will further improve coverage and geometric diversity. Recently, testing has been performed to determine GPS receiver perfor-

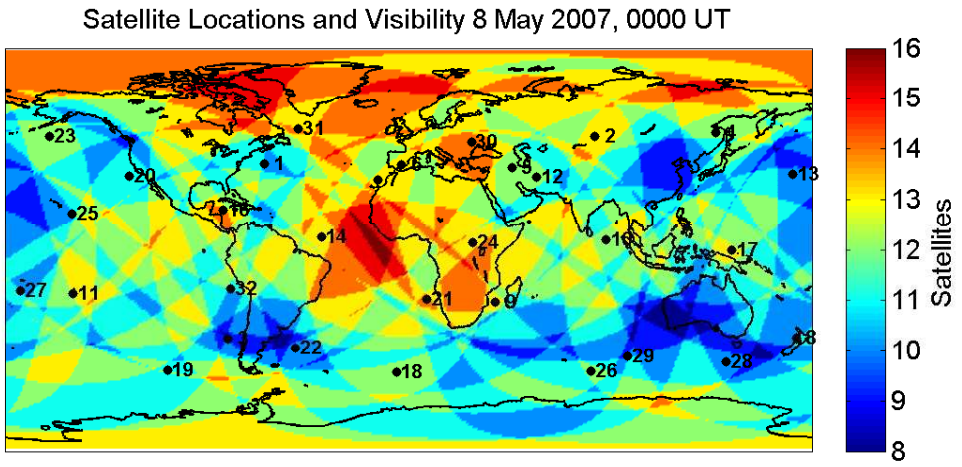

Figure 3.2: Total number of satellites visible above 0° elevation for a receiver located on the earth at 0000 UT on 8 May 2007. 8 or more satellites were in view for the entire earth, with 16 satellites in view over the tropical Atlantic Ocean. The black dots indicate the sub-satellite positions of each GPS satellite (indicated by its PRN number) at 0000 UT.

mance using a constellation of more than 31 satellites[1]. Figure 3.2 is a plot of the earth with the sub-satellite points plotted. The color scale indicates the number of satellites in view as observed from the earth. As of May 2007, any receiver on the surface of the earth that is capable of viewing from horizon to horizon would be able to track 8 – 16 satellites.

The most critical component for the navigation and timing services provided by GPS are the redundant rubidium and cesium atomic frequency standards on-board each satellite. Each Block II-R satellite flies with two rubidium standards and one cesium standard. These frequency standards have been shown to have stabilities to one part in 10^{13} over 1 - 10 days. The fundamental frequency for GPS is 10.23 MHz, although this quantity is slightly dithered to 10.229999995453

[1]http://sidt.gpsworld.com/gpssidt/GPS+Modernization/Old-Block-IIA-Sat-Reactivated-for-Tests/ArticleStandard/Article/detail/396031

Figure 3.3: Artist's conception of a GPS Block II satellite in orbit. The antennas are visible on the face of the satellite. Image © Lockheed Martin Corporation, reproduced with permission.

MHz to account for relativistic effects from reduced gravitational potential and the satellite orbital velocities (*Allan et al.*, 1998). The atomic clocks provide the frequency stability needed for the GPS signals to be transmitted synchronously from 24 or more satellites and ensures that the satellites need not be continuously monitored for frequent updates of the ephemerides. All GPS signals are commonly generated using the atomic standard on-board each satellite, and it is these signals that the receiver processes to give a measure of time and position. The precision of the frequency standard on board the GPS satellites also frees the user of expensive and bulky onboard atomic clocks.

From the frequency standard, the satellite is responsible for generating and broadcasting the two primary transmission frequencies at 1.57542 GHz (L1) and 1.2276 GHz (L2) and the ranging codes derived from pseudo-random noise se-

quences. These signals are discussed in Section 3.4, or for a more detailed account, please refer to *Spilker* (1996a). In addition to the transmitted frequencies and ranging codes, each satellite broadcasts a 50 Hz navigation message that transmits each satellite's ephemeris, a measure of GPS time, corrections for ionospheric delay for single frequency users, almanacs for all other satellites, and other various system parameters (see Section 3.5). GPS time is further discussed in Section 3.2.2.

3.2.2 Control Segment

The GPS Control Segment is responsible for the satellite health and updates to the satellite navigation message and clocks. The center of the Control Segment is the Master Control Station (MCS) located at Schriever Air Force Base near Colorado Springs, Colorado. The Control Segment is tasked with the following functions (*Francisco*, 1996):

1. Control and maintain the status, health, and configuration of the SV constellation,
2. Support the user segment with precision predictions of ephemeris and time scale calibration data. Prepare and upload the formatted navigation message data sets to the SV for subsequent metered retransmission to the user community,
3. Monitor the quality of navigation and time transfer services as provided to the end users,
4. Support system interfaces to associated services (i.e. United States Naval Observatory),

5. Manage and schedule the ground assets of the control segment.

The satellite signals are tracked from from U.S. Air Force monitor stations located around the globe (east-to-west): Ascension, Cape Canaveral, Colorado Springs, Hawaii, Kwajalein, and Diego Garcia. The original network guaranteed that every satellite was in view from one station at all times. The stations are essentially GPS receivers with cesium clocks and supporting meteorological instruments, and communications equipment to send the data to the MCS.

Each site is operated remotely and monitors the signals to determine the satellites' health. If an anomalous event is detected, the satellite is given an unhealthy code until the event is corrected or the satellite is removed from the constellation. The satellite orbits and clock biases are predicted by the MCS using data collected from the satellites at the ground stations. At least once per day, the new satellite ephemeris and clock biases are uploaded to the satellites via S-band radio links situated at Cape Canaveral, Kwajalein, Diego Garcia, and Ascension. Since 1995, six additional monitoring stations, operated by the National Geospatial-Intelligence Agency (NGA) of the DoD have been added at Washington, D.C., United Kingdom, Argentina, Ecuador, Bahrain, and Australia. The addition of these stations guarantee that each satellite is monitored by at least two stations. The additional monitoring has also improved the clock bias and satellite orbit prediction. Five more NGA-operated monitor stations will be added to the network in the near future (*Misra and Enge*, 2006).

GPS satellites do not transmit UTC time but rather a quantity known as GPS time. GPS time officially started at midnight 6 January 1980. GPS time is a count maintained by two fields–the GPS week which can range from 0 to 1023, and the GPS seconds into the week which ranges from 0 to 604,800 seconds. The

GPS week counter increases by one at midnight every Sunday at which point the GPS seconds reset for the week. Every 1024 weeks the week counter resets to zero (this last occurred on 22 August 1999). The archaic method of keeping time is simply a result of technological limitations at the time the system was created– microcomputers and large-scale integrated circuits were not available. The system cannot tolerate jumps in time, so it is not strictly coordinated with UTC time, which suffers from the addition of leap-seconds. However, it is the responsibility of the control segment to maintain the GPS second aligned to within 97 ns of the UTC second (*Francisco*, 1996). In reality, the control segment is capable of maintaining the UTC second aligned to within 10 ns (*Allan et al.*, 1998). Managed by the control segment, a parameter within the navigation message is transmitted to each user indicating the number of leap-seconds between the UTC and GPS time references. As of May 2007, GPS time leads UTC by 14 seconds.

3.2.3 User Segment

The user segment involves any user of the GPS satellite signal. At the most basic level, any device that is capable of receiving and processing GPS signals is part of the user segment. The military employs GPS for navigation of aircraft and ships, precision bombing, and time synchronization of military operations. The public and private sector utilize GPS for aviation, ship traffic, spacecraft, cell phone users, ATM and other banking transactions, the internet, power grid synchronization, in-car navigation systems, 911 services, precision timing, offshore oil platforms, and many others. The scientific community applies GPS for geodesy, plate tectonics, gravity field mapping, water vapor and climactic measurements, animal migration, and ionospheric sciences among others.

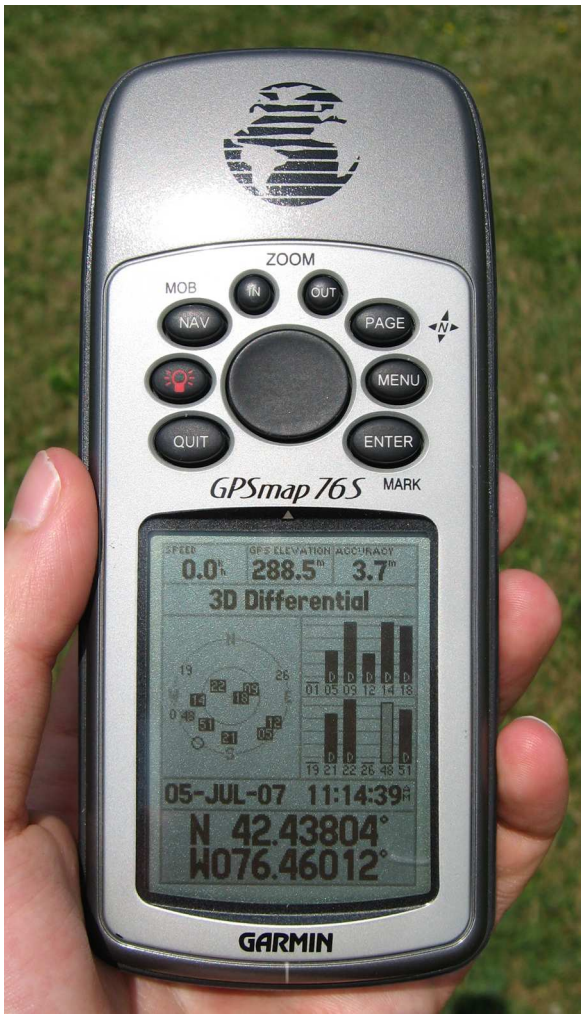

Figure 3.4: A Garmin GPSmap 76S 12-channel GPS receiver in operation. The receiver is showing the date and time, latitude, longitude, altitude, accuracy, visible satellite signal strength, and an elevation-azimuth map.

The user segment has grown significantly since the first receivers were developed in the early 1980s. The first precise receivers introduced in the mid-1980s cost over $100,000. Some receivers were only capable of tracking one satellite signal at a time. Today, receivers with even more capability only cost $5,000. Most receivers today are capable of tracking 12 or more satellites simultaneously. In 1992 the price of a hand-held GPS receiver fell below $2,000, and in 1997 the first receivers priced under $100 were available (*Misra and Enge*, 2006). In 2003, according to the U.S. Department of Commerce, the worldwide sales of GPS products and services

exceeded $16 billion. The capability of GPS to provide very precise timing and navigation capabilities coupled with the explosion of low-cost GPS receivers have made GPS a vital, free global utility.

3.3 Global Positioning System Performance Requirements

According to *Spilker and Parkinson* (1996), the original performance requirements of the Global Positioning System are as follows:

1. High-accuracy, real-time position, velocity, and time for *military users* on a variety of platforms, some of which have high dynamics; e.g. a high-performance aircraft–high accuracy translates into 10-m three-dimensional rms position accuracy or better, velocity accuracy better than 0.1 m/s,
2. Good accuracy for *civil users*–the real-time civil user accuracy objective is considered to be 100 m (at about the 95th percentile) or better in three dimensions. In the future, this accuracy may be improved by reducing or eliminating the deliberate degradation of the ranging signal,
3. Worldwide, all weather operation, 24 hours a day,
4. Resistance to intentional (jamming) or unintentional interference for all users–enhanced resistance to jamming for military users,
5. Capability for highly accurate geodetic survey to centimeter levels using radio frequency carrier measurements–capability for high-accuracy time transfer to 100 ns or better,
6. Affordable, reliable user equipment–users cannot be required to carry high-accuracy clocks; e.g., atomic frequency standards, or sophisticated arrays of

directional antennas that must be pointed at the satellites.

The performance standards today are met, and in most cases exceeded. For military users, total errors usually amount to a few meters. As of May 2001, the deliberate degradation of the civilian signal was terminated by a Presidential Decision Directive. The *GPS SPS Performance Standard* (2001) guarantees a global average positioning error of ≤ 13 meters horizontally and ≤ 22 meters vertically, and at most 36 meters horizontally and 77 meters vertically. Civilian time transfer is guaranteed to be better than 40 ns. Civil users frequently enjoy errors as small as 3-5 meters when using augmentation systems, significantly smaller than those guaranteed by *GPS SPS Performance Standard* (2001). By measuring the phase of the carrier signal from the satellite, employing differential techniques, or averaging, civilian users are now capable of precision at the sub-centimeter level. For the time being, civil users are still limited to single-frequency measurements, so the ionosphere is by far the largest contributor to the error budget.

3.3.1 Ranging Errors

The ranging errors that plague the global positioning system can be grouped into the following three classifications:

1. Satellite ephemeris and clock modeling errors,
2. Signal propagation errors due to the ionosphere and the troposphere that cause a bias in the measured range to the satellite,
3. Receiver dependent errors including noise and multipath.

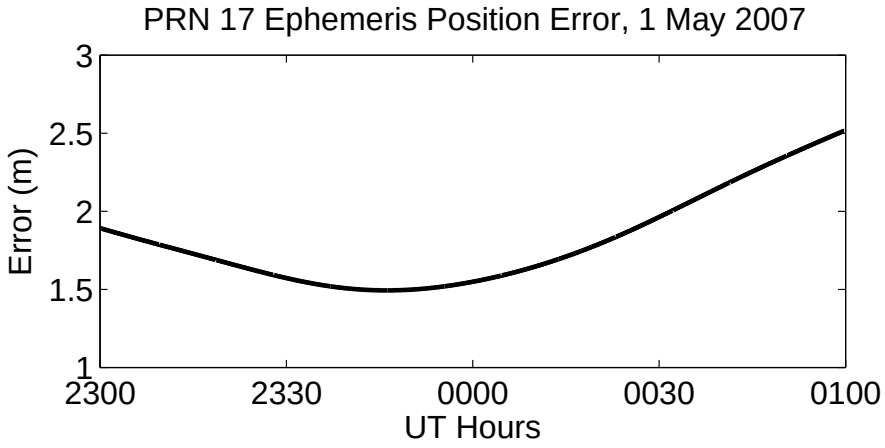

Figure 3.5: The GPS Control Segment maintains the satellite orbit and calculates the transmitted satellite ephemeris using an extended Kalman Filter from data collected at the ground stations. The plot above shows the 3-D error in calculating the satellite position using the transmitted ephemeris as compared to the 13 day latency International GNSS Service (IGS) Final Orbits, which have an accuracy better than 5 cm.

3.3.2 Satellite Ephemeris and Clock Modeling Errors

The satellite ephemeris and clock parameters are computed by the Control Segment using actual measurements from the satellites. The measurements are processed by a Kalman filter to generate the future uploads of satellite ephemeris and clock parameters. The measurements grow with time, and so more frequent updates of ephemeris and clock parameters to the satellites results in lower Control Segment errors.

By specification of *GPS SPS Performance Standard* (2001), the clock and ephemeris errors are limited to 6 meters. The metric is exceeded, however, as the stability of the satellite clocks and the frequent uploads (every two hours) by the control segment has led to errors better than 3 meters (see Figure 3.5). The ephemeris errors are typically decomposed into three orthogonal directions relative to the satellite orbit: radial, along-track, and cross-track. Fortunately, for GPS,

the smallest errors are in the radial direction, since radial errors cause the largest errors in the navigation solution (*Misra and Enge*, 2006).

3.3.3 Signal Propagation Errors

Signal propagation is often the largest source of error for both single and dual-frequency users. The Earth's neutral atmosphere changes the index of refraction for radio signals propagating through it. The electrically neutral atmosphere, known as the troposphere, is the region of the atmosphere extending to an altitude of 40 km. The following discussion on tropospheric effects is derived from *Spilker* (1996b).

About 75% of the delay caused by the neutral gasses of the atmosphere occurs in the region of the atmosphere known as the troposphere. Although 25% of the delay is attributable to the region of the atmosphere above the troposphere, the effects of the neutral atmosphere are collectively known as the tropospheric effects.

The neutral atmosphere is non-dispersive for GPS frequencies, and so the refractive index does not depend on the signal frequency. Furthermore, the phase and group velocities of the signal remain the same. The overall effect of the neutral atmosphere causes a ranging error of approximately 2.5 m for high elevation satellites to about 25 m for low elevation satellites. The tropospheric delay is a function of both the dry gas and the wet gas (water vapor), each of which are estimated using a model. About 90% of the delay is due to the dry gas, which is fortunate because the delay due to the wet gas is significantly more difficult to model. Furthermore, the tropospheric delay has a variability of about 30 cm that is attributable to the wet component. Many models exist to model the tropospheric

delay, including the Saastamoinen model and the Hopfield model. Both models determine the zenith (vertical) delay, which then must be corrected for a satellite at a given elevation using mapping functions. Even simple models, which do not account for local meteorological conditions, have modeling errors of only about 5 - 10 cm.

Incidentally, the troposphere causes a maximum signal attenuation of about 0.5 dB for low elevation satellites. Even rainfall of 100 mm/h (equivalent to about 3 in/h) only causes attenuation of about 0.01 dB/km at 2 GHz and is even less for the GPS frequencies. Rain and cloud cover do not affect reception of the GPS signals.

In addition to the neutral atmosphere, there exists a portion of the atmosphere that is a charged gas (plasma) that lies at altitudes above 100 km. This portion of the atmosphere is known as the ionosphere and is discussed in more detail in Chapter 2. The ionosphere is not nearly as predictable as the troposphere and is highly variable both spatially and temporally. The ionosphere is a dispersive medium for radio waves, which causes a phase advance but a group delay of the signal (Section 2.2). For GPS, this translates to an advance of the L1 and L2 carrier signals, but a delay in the arrival time of the pseudorandom sequences that is proportional to the Total Electron Content (TEC) of the ionosphere. A TEC Unit (TECU) is equal to $10^{16} \frac{e^-}{m^2}$ and causes a delay of about 16 cm at the GPS L1 frequency. The zenith delay of the ionosphere can amount to several tens of meters, and depends diurnally, seasonally, on solar cycle activity, solar flares, and geomagnetic variations. The delay increases for low elevation satellites. Typically the zenith delays are largest below the equatorial ionization anomaly (EIA), which occurs about $\pm 15^\circ$ from the geomagnetic equator during the local afternoon hours.

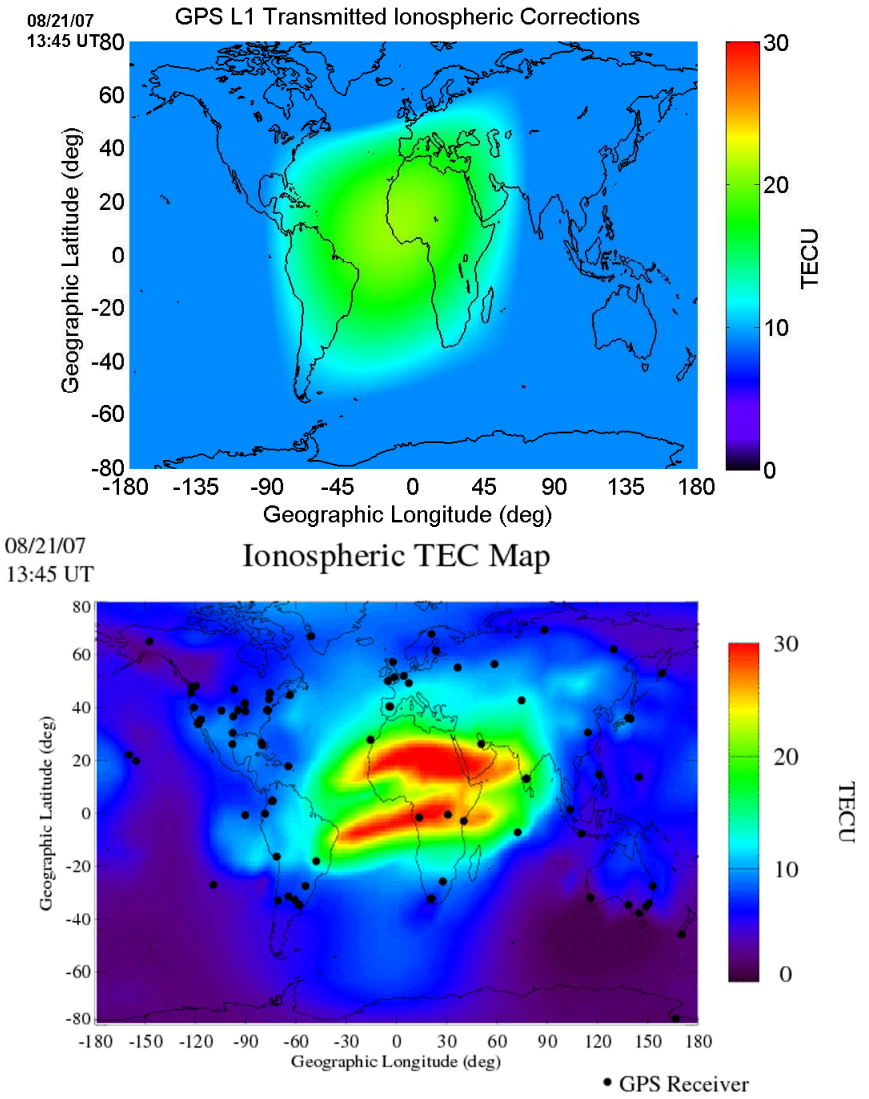

Figure 3.6: The top panel shows the vertical Total Electron Content of the ionosphere as estimated by the GPS transmitted model. The bottom panel is an estimate of the true state of the ionosphere created from world-wide dual-frequency observations of GPS satellites. The color scales are approximately equal. Notice the large discrepancies in the transmitted model versus the estimated state of the ionosphere. These discrepancies will result in ranging errors for single-frequency GPS users. Bottom panel courtesy of NASA/JPL (http://iono.jpl.nasa.gov/latest_rti_global.html).

With a dual-frequency GPS receiver, the differential arrival times of the L1 and L2 signals allow for accurate determination of the TEC between the user and satellite. Dual-frequency receivers are capable of correcting for ionospheric ranging errors. For the single-frequency user, the delay induced by the ionosphere cannot be measured. Instead, the GPS satellites transmit a coarse model of the ionosphere that is applied to attempt to remove the bulk of the ionospheric error. During solar and geomagnetic active periods, the transmitted model and corrections can have residual errors that amount to ±30 meters or more. The ionosphere is frequently the largest source of error for single frequency users. The top panel of Figure 3.6, originally shown in Chapter 2, illustrates the GPS transmitted model of the ionosphere as compared to the bottom panel, which shows the actual state of the ionosphere as calculated from dual-frequency observations.

3.3.4 Receiver Dependent Errors

To determine the range to the satellite, the receiver must make precise ranging measurements of the phase of the arriving signals. Receiver noise is any noise introduced by the physical components of the receiver including the antenna, amplifiers, cables, receiver components, as well as noise introduced by signal quantization and interference from other GPS signals. These white-noise sources challenge the receiver's ability to accurately track the incoming signal from the GPS satellite. Noise is devastating to position accuracy when the signal-to-noise ratio is particularly weak. GPS receivers are capable of measuring the carrier phase of the incoming signals to a precision of 0.5-1% of a wavelength ($\sim$ 1.9 mm at L1) and the code phase to within 0.5 m (*Misra and Enge*, 2006).

The physical placement of the receiving antenna can also contribute substan-

tially to the error budget. Multipath occurs when the GPS signal reaches the antenna via two or more paths. The antenna may receive the direct signal, and a delayed, sometimes weaker version of the same signal reflected off a nearby surface. The ranging error introduced by the reflected signal is a function of many factors, including the delay, the indirect signal's strength, and receiver design. To minimize multipath effects, one should place the receiving antenna as far from reflecting surfaces (e.g. buildings) as possible. For a single-frequency GPS L1 C/A code receiver, multipath errors can easily amount to 5 m or worse. In benign conditions multipath can be reduced to 1 m. In severe cases, multipath errors in excess of 50 meters are possible.

The ranging errors can be summarized by a single quantity known as the user range error (URE) or user equivalent range error (UERE). UERE is the root-sum-square (rss) of the 1-σ ranging errors. The UERE is strongly dependent on the user's equipment and location, primarily due to high variability of the ionosphere and multipath environments. As will be shown later in this chapter, the root-mean-square (rms) positioning errors in GPS are proportional to the UERE. Table 3.1, derived from *Leva et al.* (1996), illustrates the various errors associated with each of the segments, the magnitude of the error, and the total rms sum: the System UERE.

Although it is beyond the scope of this thesis, a brief mention of differential GPS (DGPS) is warranted because DGPS techniques can substantially reduce ranging errors. DGPS is the technique of using a well-surveyed receiver to measure the ranging errors to each satellite. The ranging errors are then transmitted as corrections to nearby GPS receivers. Using the DGPS technique, the satellite ephemeris and clock errors as well as the tropospheric and ionospheric errors are

Table 3.1: Global Positioning System System user equivalent ranging error (UERE) for L1 C/A code measurements. The UERE is the root-sum-square of all the system errors, and the positioning error is proportional to the UERE. Note that the ionospheric delay and the multipath quantities are highly dependent on the user location and equipment. The table is adapted from *Leva et al.* (1996).

Segment Source	Error Source	GPS 1σ Error (m)
Space	Satellite clock stability	3.0
	Satellite perturbations	1.0
	Other (thermal radiation, etc.)	0.5
Control	Ephemeris prediction error	3.0
User	Ionospheric delay	5.0
	Tropospheric delay	0.5
	Receiver noise	0.5
	Multipath	2.5
	Other (interchannel bias, etc.)	0.5
System UERE	Total (rss)	7.2

nearly eliminated since they are highly correlated over short distances and time periods. Multipath and receiver noise, however, are uncorrelated from receiver to receiver and cannot be corrected. DGPS offers precision to 10 - 20 cm, and carrier phase DGPS (CDGPS) offers precision to sub-centimeter levels (*Psiaki and Mohiuddin*, 2005). For a detailed discussion on DGPS techniques, please refer to *Parkinson* (1996b).

3.4 Global Positioning System Signal Structure

The Global Positioning System (GPS) transmits right-hand circularly polarized (RHCP), direct sequence-spread spectrum code division multiple access (CDMA) signals on two frequencies which are derived from the master clock frequency of 10.23 MHz. The transmission signals are at L1 (1.57542 GHz $= 154 \times 10.23$ MHz)

and L2 (1.2276 GHz = 120×10.23 MHz). Since every satellite transmits at exactly the same carrier frequency, each transmitter is uniquely identified from the others and is tracked using a preassigned pseudorandom noise (PRN) sequence. In the current GPS signal structure used by Block II, IIA, or IIR satellites, each satellite vehicle (SV) transmits a publicly available PRN sequence on L1 known as a coarse acquisition (C/A) code and an encrypted military PRN sequence on L1 and L2 known as the encrypted precise (P(Y)) code.

A PRN sequence is a quasi-random sequence of chips (a chip is a data-less bit) that has unique mathematical properties. Since the signals are random in nature, the mean and standard deviation of the correlation value between the two signals is $\frac{1}{N}$ and $\frac{1}{\sqrt{N}}$, respectively, where N is the length of the sequence. In other words, the correlation between two identical sequences will have a maximal correlation value of $20 \log_{10} \frac{1}{\sqrt{N}}$ dB higher than the correlation between two random signals of the same length (*Sarwate and Pursley*, 1980). In this case, two differing PRN sequences are said to be orthogonal to one another. The orthogonality of the PRN sequences enables the receiver to uniquely distinguish each GPS satellite despite that they are all transmitting on the same frequency.

PRN sequences also possess autocorrelation properties that are advantageous for signal tracking. When the time-shift of the autocorrelation is restricted to be only an integer chip width, the autocorrelation of a PRN sequence can only result in the following normalized values: 1, $\frac{-1}{N}$, $-\frac{1+2^{\lfloor (n+2)/2 \rfloor}}{1023}$, and $\frac{-1+2^{\lfloor (n+2)/2 \rfloor}}{1023}$ (*Sarwate and Pursley*, 1980). In other words, the autocorrelation has a maximal value of unity when the two identical PRN sequences are aligned, and when they are misaligned by more than an integer chip, the autocorrelation takes on a value that is significantly smaller. The smallness of the sidelobes and the narrow width of

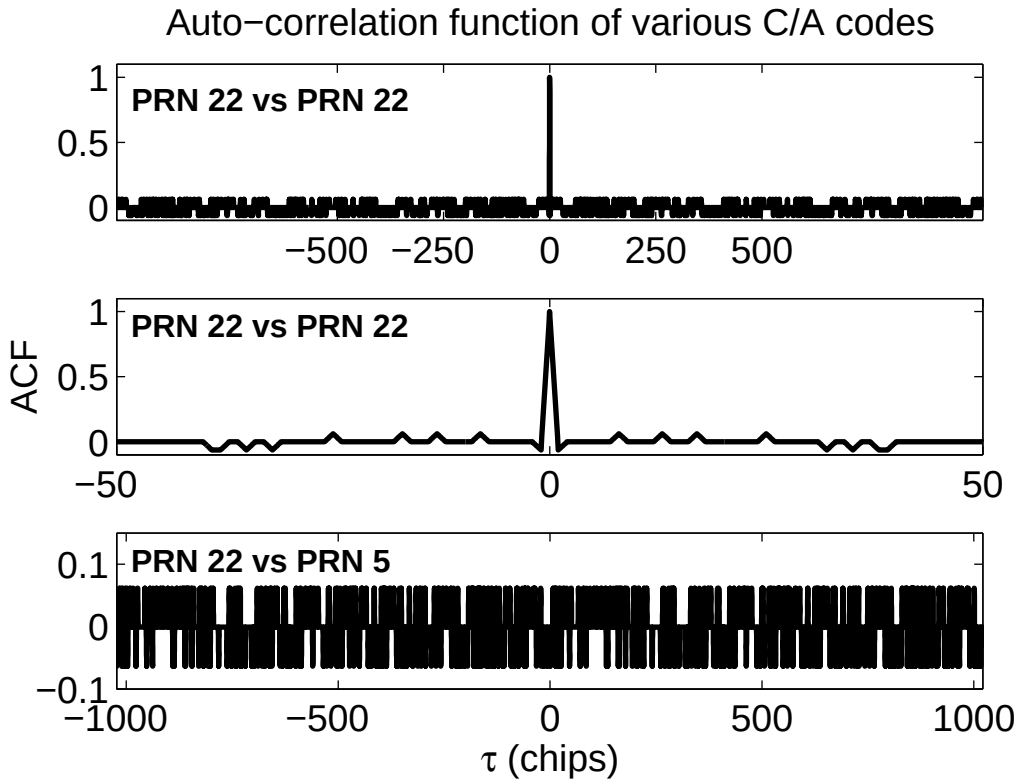

Figure 3.7: Autocorrelation function results for various GPS C/A codes (a PRN sequence of length 1023). The top two plots represent the autocorrelation function for PRN 22. Note that at 0 lag, the autocorrelation value is unity. For lags greater than one chip, the autocorrelation value is only -65/1023, 63/1023, or -1/1023, as shown in the middle plot. The bottom plot shows that the correlation between two differing C/A codes (5 and 22) is as small as the autocorrelation for lags greater than one chip. The maximal correlation between the two non-ideal, differing sequences is about -24 dB, as compared to -30 dB for two perfectly random sequences.

the aligned PRN sequences facilitates tracking of the GPS satellite signal. The autocorrelation and cross-correlation properties of the GPS L1 C/A code are illustrated in Figure 3.7. In addition to satellite tracking, the known sequence of chips in the PRN sequence allows the receiver to make unambiguous, fairly precise range measurements (known as pseudoranges) to each satellite.

Since the PRN code is a random sequence of chips, or square waves, its spectrum can be represented by a sinc function. If the duration of a chip is T_c seconds, then

the main lobe of the sinc function has a bandwidth of $\frac{1}{T_c}$ Hz. 90% of the power lies within the main lobe of the sinc function, and when the PRN sequence is applied to the carrier signal, most of the power of the carrier is spread over $\frac{1}{T_c}$ Hz. For GPS, the system designer is typically only concerned with the main lobe of the PRN sequence spectrum, and so the main lobe bandwidth of the sinc function is the bandwidth of the PRN sequence. Although the spreading of the signal by the PRN sequence spreads the power over a large bandwidth, this feature is actually a boon to the GPS signal because the spreading increases the receiver's ability to cope with narrow-band interference.

In GPS, there are primarily two PRN sequences. The first sequence, known as the coarse-acquisition or C/A code, is a sequence of 1,023 chips and is clocked at 1.023 Mega-chips per second (Mcps) (the C/A code chipping rate is $\frac{1}{10}$ the fundamental GPS frequency of 10.23 MHz). At transmission, the period of the C/A code is 1 ms, with a chip period of 977.5 ns. The C/A code is modulo-2 added to the satellite ephemeris data bits. The ephemeris data bits are transmitted at 50 bits per second (Bps), so 20 C/A codes represent a bit of ephemeris data. Finally, the entire sequence (C/A code and data bits) is transmitted only on the L1 carrier using binary phase shift keying (BPSK). In BPSK, a binary bit of 1 causes a 180° phase shift in the carrier, whereas a binary bit of 0 leaves the transmitted phase unchanged. Figure 3.8 illustrates how the in-phase component of the L1 GPS signal is created from the C/A code and the data bits.

In addition to the L1 C/A code, a PRN sequence known as the precise (P) code is transmitted in quadrature on both L1 and L2. The P code is encrypted by the United States Department of Defense (DoD) to deny access to non-authorized users using a secure anti-spoof Y code and is henceforth known as the P(Y) code.

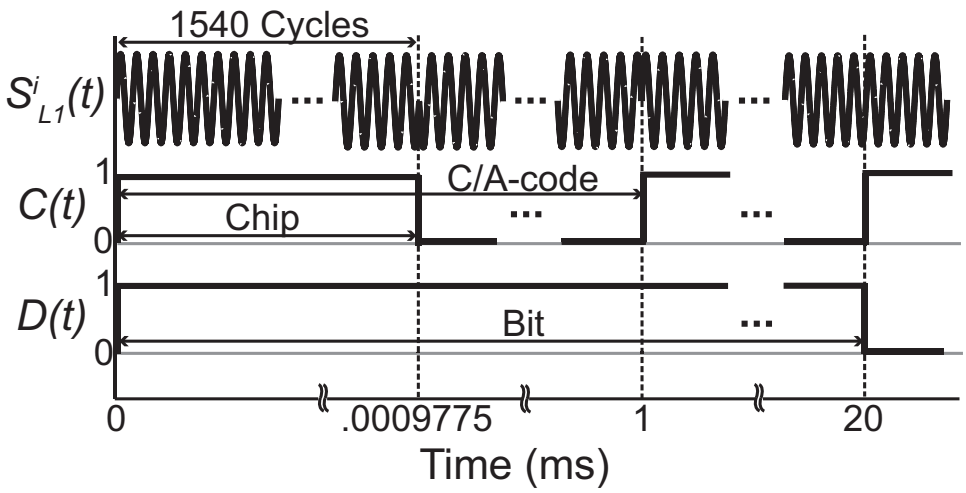

Figure 3.8: The L1 in-phase signal is transmitted as shown in the figure above. As indicated, 1540 L1 carrier cycles constitute one chip of the C/A code, $C(t)$. Each C/A code is 1023 chips long, resulting in a code of 1 ms duration. Finally, 20 C/A codes is the duration of one data bit, $D(t)$. The C/A code and data bits are modulo-2 added together, and then modulated onto the carrier resulting in the in-phase portion of the transmitted signal $S^i_{L1}(t)$.

The P(Y) code clocks at 10.23 Mcps and is reset every week, resulting in a code that is 6,187,104,000,000 chips long. The P(Y) sequence is much longer than one week in duration and so the same code is shared amongst all satellites in the the constellation, with each satellite transmitting a different portion of the code. Upon this signal the data bits described in the preceding paragraph are also encoded and the entire bit sequence (P(Y) code and data bits) are transmitted using BPSK.

Thus, for GPS, the L1 and L2 signals at transmission can be modeled as:

$$\begin{aligned} S_{L1}(t) &= \sqrt{2P_C}C(t)D(t)\cos\left(2\pi f_{L1}t+\phi\right) \\ &\quad +\sqrt{2P_1}D(t)\sin\left(2\pi f_{L1}t+\phi\right) \qquad (3.1) \\ S_{L2}(t) &= \sqrt{2P_2}P(Y(t))D(t)\sin\left(2\pi f_{L1}+\phi\right) \qquad (3.2) \end{aligned}$$

where $\sqrt{2P_C}$, $\sqrt{2P_1}$, and $\sqrt{2P_2}$ represent the signal amplitude of the C/A code and precise code on L1 and L2 respectively, $C(t)$ is the C/A code sequence, $D(t)$ is the data bits common to both L1 and L2, and $P(Y(t))$ is the encrypted precise code, also common to both L1 and L2. Finally, f_{L1} and f_{L2} are the L1 and L2 carrier frequencies, and ϕ is a phase offset.

3.5 Global Positioning System Navigation Message

The purpose of the GPS navigation message is to provide the following requirements. In addition to each requirement, a brief description of the engineering solution is included in the following list (*Spilker*, 1996c):

1. Precise satellite position at time of transmission is determined using the transmitted satellite ephemeris and a modified Kepler model in an Earth-centered-inertial coordinate frame with transformation to Earth-Centered, Earth-Fixed (ECEF) coordinates,
2. Precise satellite time at transmission is determined using satellite clock error models and relativistic correction,
3. P(Y) code acquisition from the C/A code is facilitated by a transmitted handover word (HOW) that keeps track of the number of P(Y) code 1.5-s periods that have occurred,
4. Facilitate selection and detection of satellites via a moderate accuracy almanac that gives approximate position, time, and satellite health for the entire GPS constellation,
5. Time transfer information to convert from GPS time to UTC,

6. Ionospheric correction for single-frequency users as an approximate model of the ionosphere as a function of time and user location,
7. Quality of satellite signals and data by providing a quantity known as user range accuracy (URA), which gives a quantized measure of of space vehicle accuracy.

The current GPS navigation message as transmitted on L1 is a low bit-rate message that repeats continuously. The data bits are transmitted synchronously with the C/A code epochs and the overall bit-rate is 50 bits-per-second (BPS). The message is organized into 30 bits per word, 10 words per subframe (6 s total duration), 5 subframes per frame or page (1500 bits or 30 s in duration). 25 subframes of data form a superframe (12.5 min in duration). Each word is encoded with a (24,30) Hamming code that allows for correction of single bit-errors and detection of double bit-errors. A significant portion of the data repeat every frame, and some data repeats every subframe.

The first two words of every subframe are dedicated to synchronization using a preamble, a handover word, and C/A code time ambiguity removal. The handover word changes every subframe as it contains the satellite clock time in the form of the Z-count. Words 3 - 10 of subframe 1 contain clock, health, and URA measures for that particular satellite. Words 3 - 10 of subframes 2 and 3 provide the ephemeris data to determine satellite position. Subframes 1 - 3, with the exception of the handover word, repeat the same data from frame to frame, with updates every two hours. Subframes 4 or 5, on the other hand, have 25 pages of satellite almanac data that repeat once every 12.5 minutes and are updated every 3.5 days. Subframe 4 also contains the ionospheric modeling parameters and the UTC to GPS clock correction information.

The repetitiveness of the navigation message ensures that a satellite signal can determine the antenna position within 30 seconds after acquisition. For most receivers, the almanac and UTC to GPS clock correction information is not required to navigate nor for further acquisition. Finally, the eight ionospheric modeling parameters only provide corrections to improve the navigation solution and are not required for navigation. The repetitiveness of the navigation message allows for easy determination of future bits since they repeat, and the fields that do change from subframe to subframe are easily predictable. Prior knowledge of the bits that have yet to arrive can significantly improve a receiver's ability to track weak signals. The technique of using prior knowledge of the bits to aid in acquisition and tracking is known as Assisted-GPS (AGPS) and is discussed in Chapter 6, Section 6.2.2.

3.6 Global Positioning System Signal Power

According to *IS-GPS-200D* (2004) the SV shall provide end-of-life worst-case carrier signal strength of at least -161.5 dBW for L1 P(Y) and -158.5 dBW for L1 C/A, and -164.5 dBW for L2 P(Y). The minimum received power is measured at the output of a 3 dB_i linearly polarized user receiving antenna at worst normal orientation, when the SV is above a 5-degree elevation angle. To meet the minimum power levels for L1 C/A, the satellite transmits approximately 27 watts (14.3 dBW) of power (*Aparicio et al.*, 1996).

For maximal coverage, a GPS satellite must radiate the entire disk of the earth. The signal travels approximately one earth radius (6371 km) further at the edge-of-earth (i.e. for a satellite near the horizon as viewed from the earth) rather than

at the sub-satellite point (for a satellite at zenith), which results in an additional path loss of approximately 2 dB. Radiating power into the space beyond the Earth is wasteful.

To satisfy these needs, the GPS Block IIR antenna has a peak gain of approximately 13.2 dB at a half-angle, $\alpha = 10°$. A half-angle is the angle measured from nadir. At the edge-of-earth angle of 13.85°, the gain is about 1 dB lower than the peak gain. At nadir (0° half-angle), the gain is approximately 11.3 dB. At half-angles beyond 14°, the gain rapidly rolls off so that at an angle of 20°, the gain has already dropped to 3.2 dB. A perfect antenna pattern does not exist and so there exist small side-lobes. The first side-lobe occurs at about 31° and is roughly 28 dB down from the peak gain. Finally, the gain pattern is not rotationally symmetric and also shows dependence on the position of the solar arrays (*Czopek and Shollenberger*, 1993).

From the transmitted power and the effective gain, a calculation of the incident signal power density at the surface of the earth can be performed for various half-angles, α, or their corresponding elevation on earth, θ. These results are tabulated in 3.2. As we can see, the incident power on the earth of a GPS satellite signal is extremely small. GPS is a weak-signal system. The values quoted for input power to the antenna are valid for the end of the satellite's life, so typical received powers are higher.

The biggest influence on the received signal power is the gain pattern of the receiving antenna, since the received signal power on the surface of the earth only slightly varies over all elevation angles. The orbital configuration of the GPS satellites is such that the satellites will be randomly scattered throughout elevation and azimuth for receivers located almost anywhere on earth (high latitude

Table 3.2: Received GPS L1 C/A code signal power densities as seen at the surface of the earth at various elevations (θ) or GPS satellite antenna half-angles (α). The more realistic value for atmospheric loss is from *Spilker* (1996b) and is less than indicated in *IS-GPS-200D* (2004).

	$\theta = 90^\circ$ $\alpha = 0^\circ$	$\theta \simeq 44^\circ$ $\alpha \simeq 10^\circ$	$\theta \simeq 4^\circ$ $\alpha \simeq 13.85^\circ$
Input power (dBW)	14.3	14.3	14.3
Antenna gain (dB)	11.3	13.2	12.2
Path loss (dB/m^2)	-157.1	-157.7	-159.1
Atmospheric loss (dB)	-0.5	-0.5	-0.5
Received power density (dBW/m^2)	-132.0 ($6.3{\cdot}10^{-14}$ W/m^2)	-130.7 ($8.5{\cdot}10^{-14}$ W/m^2)	-133.1 $4.9{\cdot}10^{-14}$ W/m^2

receivers will have satellites restricted to lower elevations). Most GPS antennas are patch antennas designed with an approximately hemispherical gain pattern for full sky coverage. Furthermore, the antennas are sensitive only to RHCP radiation since rejection of Left-Hand Circular Polarized (LHCP) radiation eliminates some multipath signals. A random sampling of L1 GPS antennas indicates that most patch antennas have peak gains of 3 – 5 dBic (isotropic circular) at zenith, which then falls off at lower elevation angles. The small gain at low elevations yields decreased received signal power for satellites at those elevations and helps to further reduce multipath errors. The gain pattern of the receiving antenna implies that, for most GPS receivers, satellites higher in the sky will have the highest received signal power.

Table 3.3 demonstrates the received signal power at the indicated elevation angles for a NovAtel GPS-702L[2]. The antenna has a peak gain of 5 dBic at zenith, and decreases by 13 dB to a level of -8 dBic at the horizon. The gain is even less at negative elevation angles.

[2]Novatel GPS-702L, http://www.novatel.com/Documents/Papers/GPS-702L.pdf

Table 3.3: Received GPS L1 C/A code signal power at various elevation, θ, or GPS satellite half-angles, α, using a NovAtel GPS-702L antenna. The gain of the antenna as a function of elevation influences the received signal power more than the transmitter antenna gain of the satellite.

	$\theta = 90^\circ$ $\alpha = 0^\circ$	$\theta \simeq 44^\circ$ $\alpha \simeq 10^\circ$	$\theta \simeq 4^\circ$ $\alpha \simeq 13.85^\circ$
Received power density (dBW/m^2)	-132.0	-130.7	-133.1
Antenna area $\frac{\lambda.^2}{4\pi}$ (dBm^2)	-25.4 ($2.88{\cdot}10^{-3}$ m^2)	-25.4 ($2.88{\cdot}10^{-3}$ m^2)	-25.4 ($2.88{\cdot}10^{-3}$ m^2)
Antenna gain (dBic)	5	0	-8
Received power (dBW)	-152.4 ($5.8{\cdot}10^{-16}$ W)	-156.1 ($2.5{\cdot}10^{-16}$ W)	-166.5 ($2.2{\cdot}10^{-17}$ W)

3.6.1 Signal-to-noise and Carrier-to-noise Ratios

To obtain an accurate measure of the performance of the system, one has to consider the signal power as compared to the noise power. One method to characterize the signal power to the noise power is with the signal-to-noise ratio (S/N_0). The peak signal power density S and the noise density N_0 must be known to determine the signal-to-noise ratio. For the L1 C/A code, the minimum received signal power is -158.5 dBW into a 3 dBic antenna and the one-sided bandwidth is 1.023 MHz. Thus, the peak signal power density is

$$S = -158.5 \text{ dB} + 10 \cdot \log_{10}(1.023 \cdot 10^6) = -218.6 \text{ dBW/Hz} \tag{3.3}$$

To compute the noise density, N_0, we turn to *Spilker* (1996a). The one-sided received noise density is

$$N_0 = k_B T_{eq} \text{ W/Hz} \tag{3.4}$$

where k_B is Boltzmann's constant and T_{eq} is the equivalent temperature of the receiver system.

For a receiver with a high-gain low-noise amplifier at the antenna, the equivalent noise temperature is defined as:

$$T_{eq} = T_A/L + (L-1)T_0/L + T_R \text{ K} \tag{3.5}$$

where T_A is the noise temperature of the antenna (usually referred to as sky noise), L is the line and filter loss (unit-less), T_0 is the ambient temperature of the receiver, and $T_R = T_0(F-1)$ is the receiver temperature, with F being the receiver noise figure (unit-less). All temperatures are in Kelvin. For a typical receiver, $T_A = 130$ Kelvin, $L = 1.1$, $T_0 = 290$ Kelvin, and $F = 1.259$ yielding a noise figure from Equation 3.4 of $N_0 = -205.2$ dBW/Hz.

Taking the ratio S/N_0 using 3.3:

$$\frac{S}{N_0} = -218.6 \text{ dBW/Hz} - -205.2 \text{ dBW/Hz} = -13.4 \text{ dB} \tag{3.6}$$

Hence the peak GPS signal power for any satellite lies below the noise floor of the receiver. GPS is a weak signal system. By aligning a local replica of the C/A code with the incoming signal, one can de-spread the signal energy and recover the carrier from the signal. The resulting quantity is simply the carrier power described in the first paragraph of this subsection. The best way of determining receiver performance is via the carrier-to-noise ratio (C/N_0). If the expected minimum carrier power is -158.5 dBW and the noise power density is -205.2 dBW/Hz, then

$$\frac{C}{N_0} = -158.5 \text{ dBW} - -205.2 \text{ dbW/Hz} = 46.7 \text{ dB-Hz}. \tag{3.7}$$

The C/N_0 is merely the S/N_0 plus the bandwidth of the C/A code in decibels (1.023 MHz = 60.1 dB-Hz). The carrier-to-noise ratio is typically the quantity for referring to signal power because it is a more natural quantity–a higher signal power is a larger positive number in dB, whereas for S/N_0 a higher signal power is still a negative quantity.

Many factors can influence the carrier-to-noise ratio. The *IS-GPS-200D* (2004) specifies that the minimum carrier power of -158.5 dBW accounts for a 2 dB atmospheric loss. Typically, however, the atmospheric loss at the GPS L1 frequency is much less. Thus the received power from the satellite is often higher. Furthermore, receiver performance, line losses, and antenna gain patterns can cause significant differences in the received signal power. Finally, in-band or out-of-band radio interference sources, either intentional or unintentional can adversely affect the receiver and decrease the carrier-to-noise ratio. Solar radio bursts, similar to any other interference source, act as a broad-band interference that increase the noise power spectral density, N_0, and hence decrease the carrier-to-noise ratio.

3.7 Global Positioning System Receivers

GPS receivers are not strictly devices dedicated to determining a location. GPS receivers are also employed for surveying, accurate timing, or simply to monitor the ionosphere and troposphere through measurement data. All GPS receivers, however, perform the same basic functions, and it is the goal of this section to provide the reader with an introduction to GPS receiver design.

In particular, this section is based on *Van Dierendonck* (1996) and the Zarlink Semiconductor GP2015 and GP2021 chip-sets[3]. The former reference provides many details and further references not covered here that detail the operation of various components of GPS receivers. The latter references, on the other hand, represent the components implemented in the ScintMon hardware GPS receivers (*Beach and Kintner*, 2001) and the digital storage (software) receivers at Cornell University (*Ledvina et al.*, 2003b). A brief description of these devices will be given, in particular to differentiate between software and hardware GPS receivers. Finally, an overview of the operation of a GPS receiver, including the various tracking loops, will be given.

3.7.1 System Level Requirements

At the system level, all GPS receivers are similar. The following requirements are derived from *Van Dierendonck* (1996). A GPS receiver consists of the following major components:

1. Antenna and pre-amplifier,
2. Reference oscillator,
3. Frequency synthesizer,
4. Down-conversion and digitization,
5. Signal Processing,
6. Applications Processing.

[3]Zarlink Semiconductor GP2015 GPS Receiver Front End, http://assets.zarlink.com/DS/zarlink_GP2015_MAY_05.pdf
Zarlink Semiconductor GP2021 GPS 12-Channel Correlator, http://assets.zarlink.com/DS/zarlink_GP2021_AUG_05.pdf

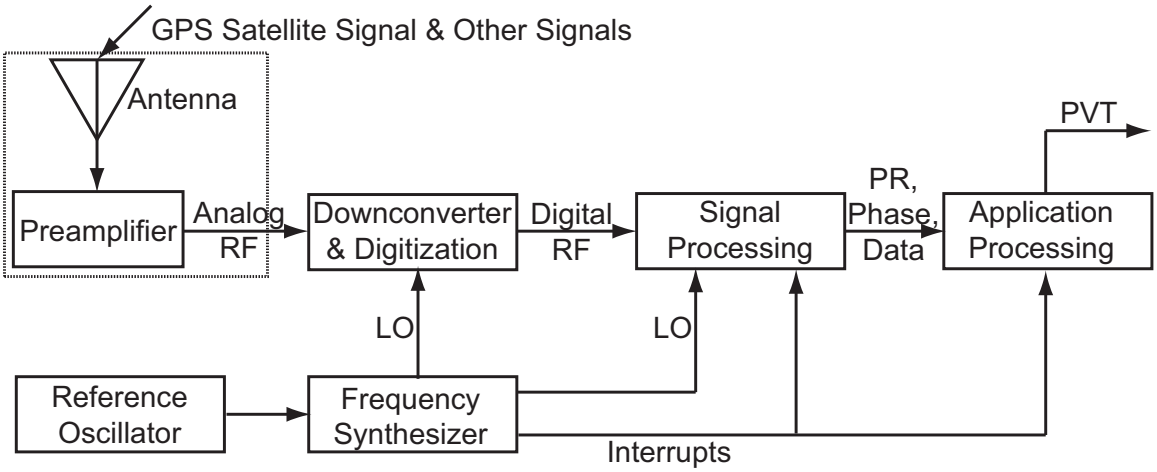

Figure 3.9: System functional block diagram of a GPS Receiver. Arrows indicate signal flow.

A functional block diagram of the major components is shown in figure 3.9. Each of the components is detailed in the appropriate sub-section below.

Antenna and Pre-amplifier

The GPS antenna and pre-amplifier are considered as one system because a pre-amplifier is typically embedded inside most GPS antennas. A GPS antenna is selected for its particular gain pattern, its ability to reject multipath and interference, phase stability, profile, size, and environmental conditions. Clearly the most important function of the antenna is to receive the GPS satellite signals. The gain pattern of the antenna determines which satellites are visible and at what received power (see Section 3.6). Furthermore, the gain pattern and polarization of the antenna determines the antenna's ability to reject multipath and in some cases interfering signals.

Interference rejection is an important consideration for any GPS receiver. Stray signals, either intentional or unintentional, and at almost any frequency, can im-

pede the receiver's ability to function properly. The antenna can provide the first line of defense by filtering out signals outside the GPS bandwidth. For interference sources within the GPS bandwidth, phased-array antennas can steer nulls in the direction of the interfering source.

Phase stability is an important consideration in surveying when using carrier phase techniques. Antennas possess a phase center which can have significant offsets from the center of the antenna that must be accounted for.

Antenna profile and size are important considerations for antennas used on aircraft, satellites, missiles, and other dynamic applications. These applications require compromises between antenna size and profile versus the desired gain patterns.

Lastly, the environmental conditions where the antenna operates determine the shape of the antenna as well as the materials used for the antenna. Care must be taken to select materials whose dielectric properties are well understood and are stable under temperature changes.

The pre-amplifier's primary responsibility is to amplify the incoming signal to a level at which the GPS receiver may process the signals. The gain of the pre-amplifier is selected to compensate for losses that precede and follow the pre-amplifier, noise and interference received through the antenna, and the desired system noise figure.

The unwanted noise and interference received through the antenna can saturate or even damage the pre-amplifier and the receiver circuitry that follows. Suppression of noise and interference from outside of the GPS bandwidth is accomplished by filtering either before, during, or after amplification. Large-signal

power interference is mitigated by the addition of diodes that provide a ground path. Lightning arrestors are sometimes used for lightning protection. Unfortunately, the pre-amplifier has no way of mitigating in-band interference.

Reference Oscillator

The oscillator in a GPS receiver is used to generate all the signals required for demodulation, detection, and measurement of the received GPS satellite signal. For many receivers, oscillator performance is not critical because the receiver has the ability to skew the clocks to compensate for oscillator drifts. There are several considerations that determine which oscillator a GPS receiver uses:

1. Size – Stable oven-controlled oscillators (OCXOs) and atomic frequency standards are rather large. Temperature-compensated crystal oscillators (TCXOs) are relatively small. Larger oscillators are less susceptible to changes in ambient temperature.
2. Power – OCXOs and atomic frequency standards consume significant power.
3. Short-term stability caused by temperature, power supply, and natural characteristics affects the ability to estimate and predict time and frequency in the receiver.
4. Long-term stability caused by natural characteristics including aging causes large offsets.
5. Sensitivity to acceleration – Vibration causes phase noise, and dynamic g forces affect the ability to estimate time and frequency in the receiver.
6. Phase noise degrades the signal-processing performance of the receiver.

Most commercial receivers employ TCXO's, while many scientific and survey grade receivers employ OCXO's. Some receivers that have onboard TCXO's allow connection of external OCXO's or atomic frequency standards. In some commercial and military applications, however, performance is critical and more accurate atomic frequency standards are required. In this case, the oscillator may cost more than the receiver.

Frequency Synthesizer

The requirements placed on the frequency synthesizer are derived requirements and are based on the *frequency plan*, which defines the receiver's IF frequencies, sampling clocks, signal processing clocks, interrupts, and so on. The frequency plan must ensure adequate rejection of mixer harmonics, LO feed-through, phase-noise, and unwanted side-bands and images. The local clocks for signal processing and interrupts comprise the receiver's time base and has many purposes including generation of local carrier and code replicas as well as determining when pseudoranges and carrier phase measurements are made.

Down-conversion and Digitization

The down-converter mixes local oscillators generated by the frequency synthesizers with the amplified RF input from the antenna to intermediate frequencies (IF) based on the frequency plan. All image frequencies, out-of-band interference, LO feed-through, and unwanted side-bands are rejected using filters in a process commonly called pre-correlation filtering. The bandwidth is chosen to avoid aliasing in a sampling receiver, and is a trade-off against correlation loss caused by filtering. Amplifiers and automatic gain control (AGC) may be included to bring the

amplitude of the signal to workable levels for digitization and subsequent signal processing.

Signal Processing

The signal processing portion of the receiver is to provide the GPS measurements and system data from the satellites to perform navigation or for other applications. Signal processing is the heart of a GPS receiver and, computationally, represents the highest workload. Signal processing performs the following functions:

1. Splitting the signal-plus-noise into multiple signal-processing channels for signal-processing of multiple satellites in parallel,
2. Generating the reference PRN codes of the signals,
3. Acquiring satellite signals,
4. Tracking the code and carrier of the satellite signals,
5. Demodulating the system data from the satellite signals,
6. Extracting code phase (pseudorange) measurements from the PRN code of the satellite signals,
7. Extracting carrier frequency and carrier phase measurements from the carrier of the satellite signals,
8. Extracting signal-to-noise ratio (S/N_0) information from the satellite signals,
9. Estimating a relationship to GPS time.

Applications Processing

Applications processing takes the necessary measurements and data from the signal processing unit to perform one or more of a variety of GPS applications. Although most GPS receivers are used for positioning, the applications processing may perform a multitude of applications, including:

1. Time and frequency transfer,
2. Static and kinematic surveying,
3. Ionospheric TEC and amplitude and phase scintillation monitoring,
4. Differential GPS reference station receivers,
5. GPS satellite signal integrity monitoring.

All these applications require the same measurements from the signal processing unit, but each application requires different bandwidth and accuracy requirements. The differing requirements typically modify the frequency synthesizer, down-conversion and digitization and signal processing functions substantially. Since special processing is required for dedicated applications, not all GPS units are capable of fulfilling each application.

3.7.2 Hardware Receiver Implementation

Historically, most receivers in the GPS community are hardware GPS receivers. That is, all the functions of the frequency synthesizer, down-conversion and digitization, and signal processing are done on non-programmable, dedicated chip-sets

known as Application Specific Integrated Circuits (ASIC). In particular, a hardware receiver baseband mixes the incoming signal and strips it of its PRN sequence using hardware correlators or fast-fourier transform (FFT) techniques. Dedicated hardware limits the system to the functional ability of the ASICs that are often designed for one application. For example, with a dedicated chip-set, it would not be possible to process new GPS satellite signals when they are added to the system. Computationally, however, the dedicated chip-sets are extremely fast and efficient.

Some popular examples of hardware receivers are the NASA/JPL Blackjack receiver, the Department of Defense Miniature Airborne GPS Receiver (MAGR), and most of the commercial and scientific receivers marketed by NovAtel, Magellan, and Garmin. Furthermore, most cell phones that are capable of Enhanced-911 (E911) location services employ hardware GPS receivers because of the low power consumption and low cost of the devices (Todd Humphreys, personal communication).

The Cornell University ScintMon 12-Channel L1 C/A code GPS Receiver is the main work-horse for ionospheric research at Cornell University (*Beach and Kintner*, 2001). The ScintMon is originally based on an Industry Standard Architecture (ISA) bus device that operated in a PC and was manufactured by Plessey (which has since been acquired by Zarlink). The current version of the ScintMon receiver has been modified by Aerospace Innovations to operate on a PC via the Peripheral Component Interconnect (PCI) bus. The receiver is capable of processing up to 12 L1 C/A code and WAAS satellite signals simultaneously, and produces 50 Hz amplitude and 10 Hz phase measurements, primarily for ionospheric scintillation monitoring. The host PC also performs the higher-level functions including

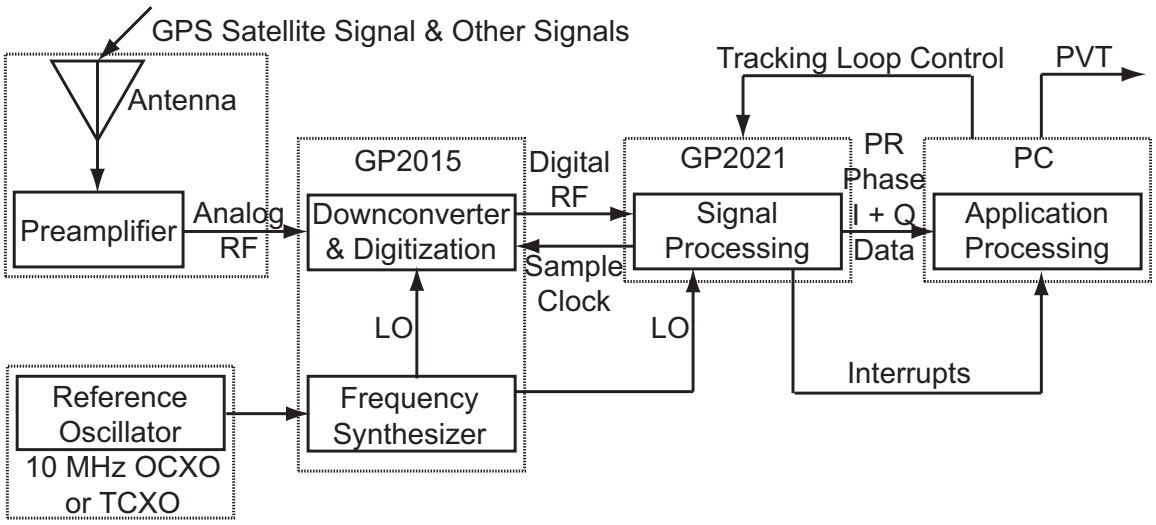

Figure 3.10: System functional block diagram of the Cornell University ScintMon 12-Channel L1 C/A code GPS Receiver. The ScintMon is a hardware GPS Receiver.

determining position, velocity, and time, as well as processing the tracking loop algorithms, and writing relevant data to disk. Figure 3.10 illustrates how the basic functional components of the Cornell University ScintMon receiver are tasked to various chip-sets and processors.

An Impact Power MAM35[4] mounted on a 50 cm by 50 cm, 3 mm thick steel or aluminum ground plane is typically the antenna of choice for use with the ScintMon. This antenna exhibits excellent signal quality and can be used with RG-58 coaxial cable up to 10 meters in length. The antenna has a peak antenna gain of 4 dBic at zenith, with only 8 dB of roll-off, and has a built-in 35 dB low-noise amplifier with a noise figure of only 2 dB. The antenna has a bandwidth of approximately 20 MHz centered around L1. The antenna consumes at most 22 mA of current at 4.5 – 6.0 volts, has dimensions of 48 mm x 40 mm x 13 mm, and is waterproof.

[4]Impact Power High Gain/High Performance GPS Antenna http://images.radcity.net/6003/889974.pdf

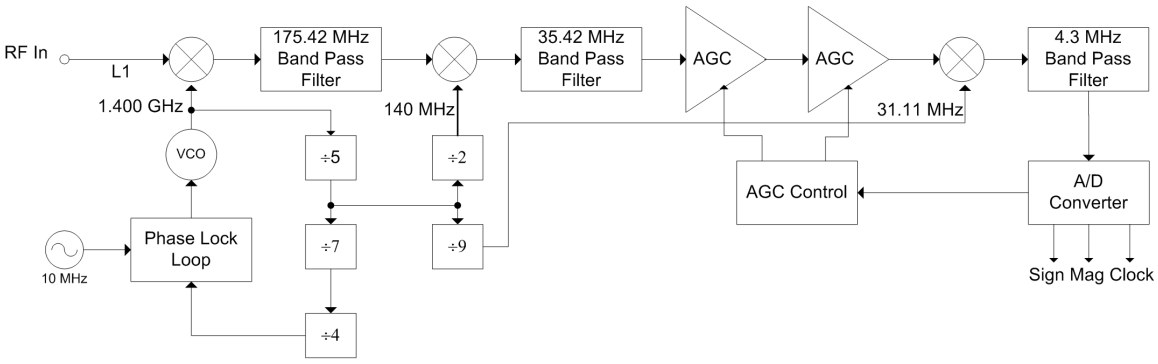

Figure 3.11: Simplified block diagram of the Zarlink GP2015. The GP2015 down-converts the GPS L1 signal from 1.57542 GHz to a center frequency of 1.409 MHz and digitizes it with a sampling rate of 5.714 MHz.

The signal from the antenna is digitized by the Zarlink GP2015. In the case of the Zarlink GP2015, it contains an on-chip synthesizer, mixers, AGC, and quantizer. From a pre-amplified L1 C/A code signal, the signal is reduced from a 1.57542 GHz center frequency to a 4.309 MHz center frequency signal using triple down-conversion. On the ScintMon implementation, the GP2015 utilizes a 10 MHz reference from either an on-board TCXO or an external OCXO. A phase-lock loop (PLL) internal to the GP2015 is the frequency synthesizer that creates all the signals for down-conversion and sampling. The 1.57542 GHz incoming signal is first mixed with a 1.400 GHz signal down to 175.42 MHz, which is then filtered using a band-pass filter. The 175.42 MHz signal is then mixed down to 35.42 MHz using a 140 MHz LO, and is band-passed filtered using a high-quality surface acoustic wave (SAW) filter with a 1.9 MHz bandwidth at 1 dB attenuation. The signal is then passed through two AGC's before mixing down to 4.309 MHz using a 31.1 MHz LO. An on-board band-pass filter provides filtering centered on 4.309 MHz. These functions are summarized in Figure 3.11.

The 4.309 MHz signal is then passed through a 2-bit quantizer which provides the sign and magnitude bits. The magnitude data controls the automatic gain

control loop such that on average the magnitude bit is set high 30% of the time. The quantized signal is then sampled using an external clock, typically at 40/7 (5.714) MHz, provided by an external source (e.g. the GP2021 Correlator[5]). The sampling scheme is known as an aliasing down-conversion and brings the center frequency from 4.309 MHz to 1.405 MHz. The signal is now ready to be processed by the Zarlink GP2021 correlator.

The Zarlink GP2021 is a 12-channel digital baseband processor that correlates the incoming signal from the satellites with locally generated replicas at the appropriate code phase and Doppler shift. The GP2021 is the portion of the GPS receiver that performs all the fast operations. Specialized hardware is required to correlate the incoming signal at 5.714 MS/s with the locally generated replicas in real-time over all 12 channels.

Illustrated in Figure 3.12, each channel of the GP2021 multiplies the input signal with a locally generated cosine and sine replica of the carrier at the appropriate intermediate frequency and Doppler shift to obtain the base-band in-phase and quadrature components. The in-phase and quadrature components of the base-band signal are then divided into two more signal paths. One signal path of the base-band in-phase and quadrature components is correlated with a local replica of the C/A code, and the other path is correlated with a local replica of an early-minus-late replica of the C/A code. Each of the four outputs are then integrated over one millisecond to obtain the in-phase and quadrature values. The four correlator values are primarily used in the signal tracking algorithms, which in turn provide the future carrier and code phase estimates. The correlators and signal tracking algorithms are further explored in Section 3.7.4.

[5]The 5.714 MHz signal is derived from a 40.0 MHz signal originally created on the GP2015 chip-set.

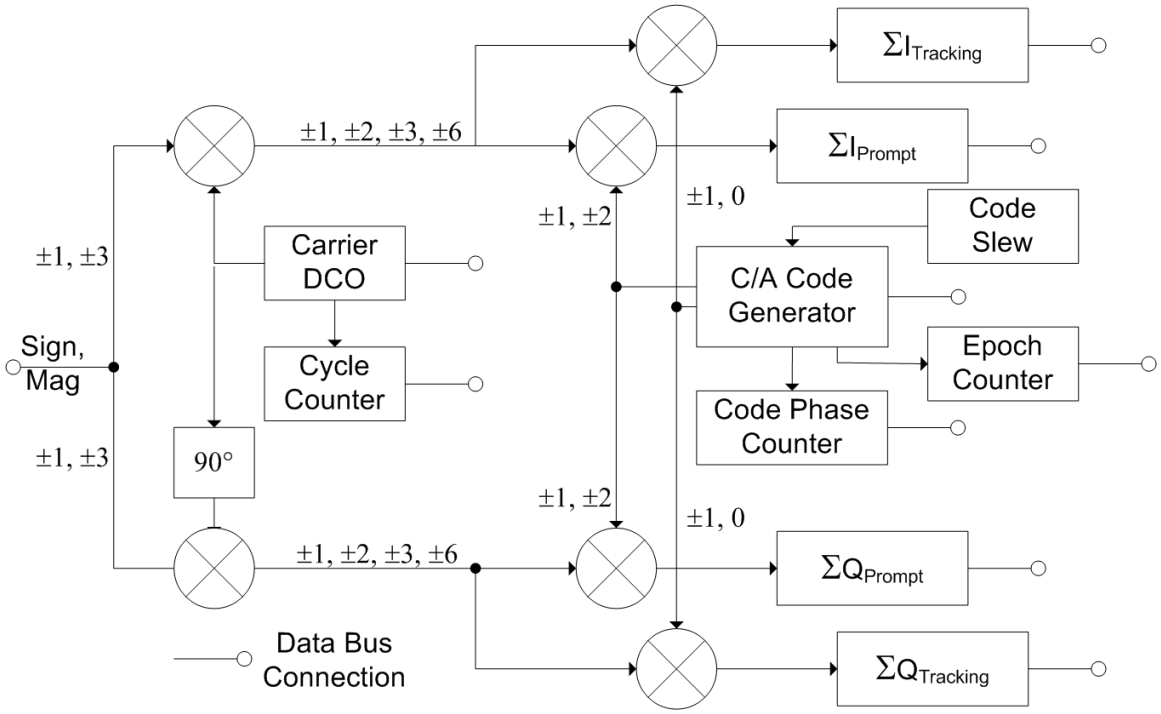

Figure 3.12: Simplified block diagram of one tracking channel of the Zarlink GP2021. The digitized, intermediate frequency GPS L1 signal from the GP2021 is base-band mixed and code stripped of the signal. From the carrier cycle counter and code phase counters the GPS observables (carrier phase and pseudorange) are determined.

The unit primarily outputs measures of the correlation values, pseudorange, and carrier phase, which are relayed through an on-chip peripheral device to the external applications processor and memory modules. The GP2021 primarily takes as input the carrier frequencies and code chipping rates from the tracking loop algorithms calculated in the applications processor. Secondary inputs include which satellite to track, and code slew and carrier slew values for initial acquisition of satellites signals. Acquisition, tracking, data demodulation, and pseudorange measurement will be further discussed in Section 3.7.4.

The unit also has an on-chip clock generator, which creates the clocks for sampling the incoming signal from the GP2015, as well as clocks for generating

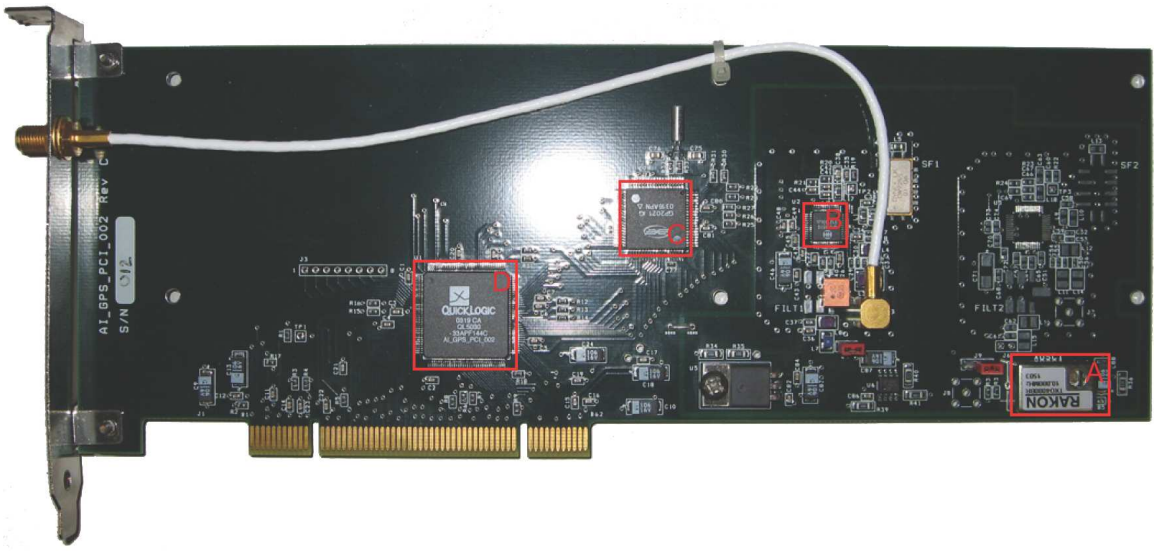

Figure 3.13: Photograph of a Cornell University ScintMon receiver. The white coaxial cable brings the signal from the antenna to the GP2015. Part A is the on-board TCXO, part B is the Zarlink GP2015 chip-set, surrounded by the appropriate filters and electronics, part C is the Zarlink GP2021 chip-set, and part D is an FPGA used to facilitate transfer of data between the correlator and the computer's PCI bus.

the C/A code sequences and appropriate carrier frequencies for demodulation. The on-chip clock also generates the measurement interrupts that are used for determining code phase (pseudorange) and carrier phase. The interrupts also drive the applications processor so that data from the correlator is processed in a timely manner.

Finally, a PC running Linux with a kernel modified for real-time using the Real Time Application Interface (RTAI) for Linux from the Dipartimento di Ingegneria Aerospaziale - Politecnico di Milano (DIAPM). The PC is responsible for processing the slow functions of the GPS receiver. Every millisecond the accumulated values from the correlators are read and stored on the PC. Once a 20 ms bit period has elapsed, the accumulated correlation values are processed in the tracking loop algorithms to determine Doppler shift and code phase, which are then sent

back to the GP2021 for the next twenty milliseconds of signal processing. The 20 ms accumulated correlation values also determine the signal amplitude, which is used for ionospheric monitoring. The correlation values are also used to determine the bit sequences, and in combination with the pseudoranges and carrier phases, the navigation solution is calculated every 100 ms.

3.7.3 Software Receivers

In a software receiver approach, the signal is digitized as early as possible using hardware, and then digital signal processing is performed in real-time using software techniques on re-programmable Field-Programmable Gate Arrays (FPGA)[6], Digital Signal Processors (DSP), or personal computers (PC). The main advantage that a software receiver possesses is flexibility. For example, the addition of new satellite signals merely requires that the chip be reprogrammed to recognize and process the signals. With the rapid growth of GPS, the lifetime of commercial components has fallen to less than two years (*Woon et al.*, 2006), and software approaches are relatively inexpensive to implement because a programming change typically requires less time and money to implement.

The main disadvantage of a software implementation, however, is computational speed and power. A PC, for example, has many additional capabilities that are not required for processing GPS signals, which limits the processor's speed and requires additional power to operate. On the other hand, a DSP may have too little computational power, or an FPGA may have too few gates to accommodate future additions. The bottom line is that a software receiver has great flexibility,

[6]An FPGA is a software approach because they are programmed using VHSIC Hardware Description Language (VHDL), but they offer the performance of hardware ASICs because they are application-specific designs.

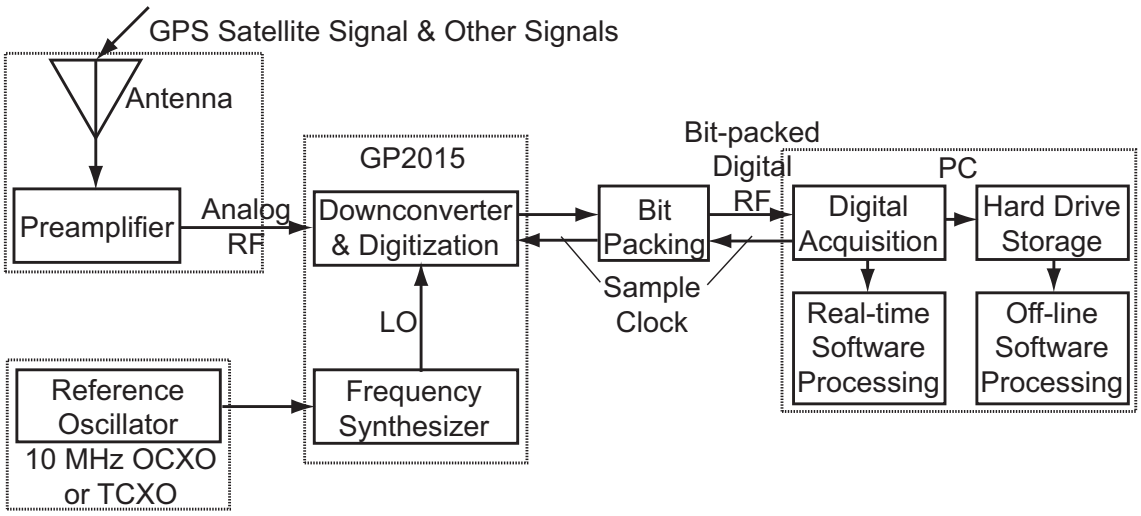

Figure 3.14: System functional block diagram of an L1 C/A code digital software receiver implementation at Cornell University. When attached to a PC, the user has the ability to store the digitized signal either on a hard disk for later signal processing or to process the signal in real-time.

but the processor must be chosen carefully to meet the required criteria.

In recent years, several companies have begun to produce embedded software receivers to support the mobile electronics market including NordNav, SiRF, and Philips (*Woon et al.*, 2006). NordNav and IfEN also offer commercial front-end units for research and development purposes, which bring digitized signals into a PC for further processing via the Universal Serial Bus (USB). The University of Colorado has also published the details to build a front-end under an OpenGPS forum.

At Cornell University, several software receivers have been developed in response to support a growing need for quick access to new signals. All the software receivers rely on the same unit for acquiring the GPS signals, which are subsequently processed in various software receivers. An overview of the various software schemes is given in Figure 3.14.

Figure 3.15: Photograph of a Cornell University L1-only bit-grabber. The card towards the front of the photograph contains the GP2015 and RF-front end components. The circuitry towards the rear bit-packs the 2-bit signal from the GP2015 chip-set into 32-bit words for transfer to the PC and storage media.

A requirement of a software receiver is that the incoming GPS signal must be digitized as closely to the antenna as possible. At Cornell University, all signals are processed using a Zarlink GP2015 front-end. The RF section of the chip-set (the frequency synthesizer, down-conversion, and digitization) is exactly the same as detailed above for the hardware Cornell ScintMon receiver. The data must then be brought into a computer for software processing.

The digitized signal as it exits the GP2015 is a two-bit serial representation of the incoming intermediate frequency analog signal sampled by the external sampling clock. The signal processing is done on-board a 32-bit PC, so it is an inefficient use of system resources to sample the two-bit signal at 5.714 MS/sec. Instead, application-specific hardware was built to take the two-bit signal and pack it into 32-bit words (this function is defined as Bit Packing in Figure 3.14). Thus, each 32-bit word represents 16 samples of 2-bit data. The 32-bit words are then

read into a PC using a National Instruments PCI-6533 digital acquisition unit[7]. The advantage of importing the bit-packed data is that the digital acquisition unit only has to sample the bit-packed words at 1/16 the original sampling rate, or about 357 kS/s. The sampled data accumulates in memory at a rate of about 1.4 megabytes per second, and can either be immediately processed in a real-time software receiver utilizing bit-wise processing techniques (e.g. *Ledvina et al.* (2003a), *Ledvina et al.* (2003b), *Ledvina et al.* (2004a), *Ledvina et al.* (2004b), *Ledvina et al.* (2005), and *Ledvina et al.* (2006)), or stored on a hard disk for later processing with specialized algorithms written in Matlab or the C/C++ programming language. A photograph of the digital acquisition unit is shown in Figure 3.15.

The Zarlink GP2015 front-end can also be utilized with other software receivers. At Cornell University, significant effort by several individuals at Cornell University including Todd Humphreys, Brent Ledvina, and many Master of Engineering students has been made to develop a DSP software receiver that is hoped will eventually replace the need for the ScintMon hardware receiver (*Humphreys et al.*, 2006).

3.7.4 Receiver Operation

A Global Positioning System receiver must process the signals arriving from the satellites to determine the users position, velocity, and time. Many schemes exist for processing the received signals, but in general they are all based upon the same principles.

The GPS satellite signal, when it arrives at the receiver's antenna, has under-

[7]http://sine.ni.com/nips/cds/view/p/lang/en/nid/13508

gone code and carrier phase shifts due to different propagation path lengths and the dispersive effects of the ionosphere, a degradation in signal power, and has a Doppler shift due to the relative motion of the satellite and receiver.

When a receiver is powered on for the first time, it may have no prior knowledge of its location or of the satellite visibility environment. This scenario is known as a cold start. In a cold start, the receiver must conduct a thorough search to find the visible GPS satellite signals. This search is known as initial acquisition. A warm start occurs when the receiver has prior knowledge of its location, an estimate of the current time, and a GPS satellite almanac to determine the expected Doppler shifts from the satellites, thus hastening satellite acquisition.

During a cold start, the receiver must search through frequency space to find a GPS signal. For an earth-bound, stationary receiver, the Doppler frequency at L1 can span ± 5600 Hz (*Spilker*, 1996a). In increments of at most 500 Hz of Doppler offset, the receiver strips the incoming signal of its carrier in a process known as carrier wipeoff. The received signal is actually demodulated using both an in-phase (cosine) and quadrature (sine) signal, primarily enabling the receiver to demodulate the signal without knowledge of the carrier phase. The data bit modulation and the C/A code modulation are still present on the signal, however.

At this point the receiver dwells at each Doppler offset and attempts to strip the incoming signal of the C/A code by slewing the start time of a locally replicated C/A code. Recall from Section 3.4 that the properties of the C/A code essentially give a null autocorrelation for lags greater than a chip, and when the C/A codes are properly aligned, a strong correlation peak occurs. Slewing the start time of the C/A code over all 1023 chips in half-chip increments allows the receiver to test if the signal is present at the current Doppler offset.

The correlations typically range from 1 ms (the duration of a C/A code) and rarely exceed 20 ms (the duration of a bit period). Longer correlation allows for more certain acquisition of the signal since the integrated signal will be detected from the random noise. There is a risk, however, that a bit transition will occur over the interval, resulting in a near-zero or diminished correlation due to the phase reversal of the signal. The magnitude of the correlation is used as a test metric to determine if a satellite is found. A probability of acquisition failure can be assigned as a metric for deciding whether a real satellite signal has been found, and is a function of the carrier-to-noise ratio, the integration time, and the number of correlations conducted (*Misra and Enge*, 2006). If the satellite signal is not found, then the receiver attempts to find the C/A code again at the next Doppler offset. The process repeats until a signal is found or all Doppler offsets are exhausted. If all Doppler offsets and C/A code phases are exhausted, the receiver repeats the process for a different satellite.

When a C/A code correlation peak is found at a particular Doppler offset and code phase, the receiver begins signal tracking. Each GPS receiver utilizes two tracking loops working in tandem to track the signal and provide necessary measurements. The satellites are in motion, so the Doppler shift and the range to the satellites is continuously changing.

A control system known as a delay lock loop (DLL) is used to maintain the local replica of the pseudorandom (C/A) code aligned with the C/A code from the satellite. The DLL uses as measurements the in-phase and quadrature correlation values, in conjunction with measures of slightly advanced and delayed in-phase and quadrature correlation values from additional correlators to align the phase of the locally generated C/A code with the incoming signal. The DLL allows the

receiver to accurately measure the C/A code phase to better than 3 meters. The measure of the C/A code phase is known as a pseudorange because it represents an unambiguous but noisy estimate of the real range to satellite. In addition, the phase transitions of the signal represent the bit-transitions of the transmitted navigation message, which is then processed to determine the time, satellite positions, and other relevant parameters.

In addition to the DLL, a dual Costas loop in conjunction with a phase lock loop (PLL) or frequency lock loop (FLL) tracks the phase of the incoming carrier signal. A dual Costas loop is implemented because it is insensitive to the carrier phase transitions induced by the navigation data bits. Both the FLL and PLL filter past measurements of Doppler shift and current and past phase errors to align the phase of the locally generated carrier signal for accurate carrier demodulation. Doppler frequency estimates obtained from the PLL or FLL aid the DLL in tracking the pseudorandom code of the satellite, thus decreasing the noise bandwidth of the DLL and improving its performance. The PLL or FLL is the heart of the GPS receiver and is the most susceptible component of the receiver in low carrier-to-noise ratio environments. The susceptibility of the PLL to low carrier-to-noise ratios will be examined in Chapter 4, Section 4.2.3. A thorough review of GPS tracking loops is provided by *Van Dierendonck* (1996) and *Ward* (1996b).

A receiver with a PLL is capable of measuring phase to about 1% of a cycle, but the ambiguity of the carrier signal does not allow the receiver to easily determine the range to the satellite. Differences in carrier phase measurements, however, are extremely accurate. In a technique known as carrier-smoothing, noisy pseudorange measurements are smoothed with differenced carrier phase measurements in a recursive filter. The precision of differenced carrier phase measurements

in conjunction with code-phase measurements allows dual frequency receivers to determine TEC to within a few TECU. Finally, carrier-phase measurements can directly and precisely solve for a receiver's velocity.

3.7.5 Navigation Solution and Dilution of Precision

A GPS receiver calculates a navigation solution (or position fix) using a technique known as trilateration. Trilateration requires that the transmitter and receiver clocks be synchronized to make precise ranging measurements between each transmitter and the receiver. Furthermore, the position of each transmitter must be accurately known. As detailed in Section 3.2.2, the GPS Control Segment takes great care to precisely determine the orbit of the satellites and to provide satellite clock synchronization corrections. The basic ephemeris and satellite clock corrections needed by a receiver to determine a satellite's position are transmitted every thirty seconds on the navigation message.

From the pseudorange measurements, a code range navigation solution is obtained by determining the coordinates of the GPS antenna attached to the receiver. The GPS receiver performs all navigation solutions in an Earth-Centered Earth-Fixed (ECEF) cartesian coordinate frame. The navigation solution is then rotated into a datum known as the WGS-84 ellipsoid model of the Earth to obtain latitude, longitude, and altitude.

The actual range from satellite k at time t to the receiver is ρ^k. Both the satellite and receiver clocks are prone to drift that causes an offset in the pseudorange. The satellite clock drift causes a correction to the pseudorange, in meters, of $c\delta_s^k(t)$, and the receiver clock causes a drift, in meters, of $c\delta_r(t)$, where c is the speed of light

(299,792,458 m/s). Thus, the pseudorange to each satellite is denoted as

$$P^k(t) = \rho^k(t) + c\delta_s^k(t) - c\delta_r(t) \text{ m} \tag{3.8}$$

The range to each satellite, $\rho^k(t)$, is related to the ECEF coordinates of the antenna $\mathbf{r}_r(t) = [x_r(t)\ y_r(t)\ z_r(t)]$ and to the ECEF coordinates of each satellite $\mathbf{r}_s^k(t) = \left[x^k(t)\ y^k(t)\ z^k(t)\right]$:

$$\rho^k(t) = |\mathbf{r}_s^k(t) - \mathbf{r}_r(t)| \text{ m} \tag{3.9}$$

Equation 3.9 is substituted into Equation 3.8 and rearranging terms yields:

$$P^k(t) - c\delta_s^k(t) = |\mathbf{r}_s^k(t) - \mathbf{r}_r(t)| - c\delta_r(t) \text{ m} \tag{3.10}$$

The satellite clock drift is well modeled by the control segment and is treated as a known quantity. If receiver clock drift did not exist, then three independent pseudoranges at time t would be required to solve for the three unknown coordinates of the receiver $\mathbf{r}_r(t)$. The receiver clock drift, however, is not a known quantity and varies with time.

The clock on the receiver is typically a low cost oscillator that has a tendency of drifting. The drift causes the receiver clock to have an error, which contributes to the error in the pseudorange measurements. Furthermore, receiver clock errors contribute to erroneous satellite positions, further causing errors in the navigation solution. The receiver clock error must be minimized. Fortunately, accounting for the receiver clock offset requires only one further observation to a satellite. The

minimum number of pseudorange measurements required for navigation is four–one for each of the three cartesian coordinates plus time.

The Newton-Raphson technique solves Equation 3.10 iteratively. The navigation solution is calculated at a single point in time, so the time dependence is no longer of concern and will be dropped from the equations. The current best guess to the navigation solution ($\mathbf{r}_{r_i} = [x_{r_i}\ y_{r_i}\ z_{r_i}]^\top$) is related to the next best guess of the navigation solution $\mathbf{r}_{r_{i+1}}$ via[8]:

$$\begin{bmatrix} \mathbf{r}_{r_{i+1}} \\ c\delta_{r_{i+1}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{r_i} \\ c\delta_{r_i} \end{bmatrix} + \begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix} \tag{3.11}$$

where $\mathbf{\Delta r} = [\Delta x\ \Delta y\ \Delta z]^\top$.

The current solution, $\mathbf{r}_{r_i}$, yields an initial range to each satellite:

$$\rho_i^k = ||\mathbf{r}_s^k - \mathbf{r}_{r_i}|| \tag{3.12}$$

To determine $\mathbf{r}_{r_{i+1}}$, $\mathbf{\Delta r}$ must be determined by using a linear expansion (first-order Taylor series) of ρ:

$$\rho^k \simeq \rho^k\big|_{\mathbf{r}_{r_i}} + \left.\frac{\partial}{\partial \rho^k}\right|_{\mathbf{r}_{r_i}} \mathbf{\Delta r} \tag{3.13}$$

Equation 3.13 is substituted into Equation 3.10 to yield:

$$P^k - c\delta_s^k - \rho^k\big|_{\mathbf{r}_{r_i}} + c\delta_{r_i} = \left.\frac{\partial}{\partial \rho^k}\right|_{\mathbf{r}_{r_i}} \mathbf{\Delta r} - \Delta c\delta_r \tag{3.14}$$

[8]For $i = 0$, guessing $\mathbf{r}_{r_0} = [0\ 0\ 0]^\top$ and $c\delta_{r_0} = 0$ suffices.

where the current estimate of the clock offset $c\delta_{r_i}$ is known and and is added to the left, and the update to the clock offset $\Delta c\delta_r$ is subtracted from the right. Setting

$$\mathbf{\Delta p} = \begin{bmatrix} P^1 - c\delta_s^1 - \rho^1\big|_{\mathbf{r}_{r_i}} + c\delta_{r_i} \\ P^2 - c\delta_s^2 - \rho^2\big|_{\mathbf{r}_{r_i}} + c\delta_{r_i} \\ \vdots \\ P^k - c\delta_s^k - \rho^k\big|_{\mathbf{r}_{r_i}} + c\delta_{r_i} \end{bmatrix} \tag{3.15}$$

$$\mathbf{G} = \begin{bmatrix} \frac{\partial}{\partial \rho^1}\big|_{\mathbf{r}_{r_i}} & -1 \\ \frac{\partial}{\partial \rho^2}\big|_{\mathbf{r}_{r_i}} & -1 \\ \vdots & \vdots \\ \frac{\partial}{\partial \rho^k}\big|_{\mathbf{r}_{r_i}} & -1 \end{bmatrix} \tag{3.16}$$

allows Equation 3.14 to be rewritten as:

$$\mathbf{\Delta p} = \mathbf{G} \begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix} \tag{3.17}$$

$\begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix}$ is easily solved using a least-squares technique:

$$\begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix} = \left(\mathbf{G}^\top \mathbf{G}\right)^{-1} \mathbf{G}^\top \mathbf{\Delta p} \tag{3.18}$$

and inserted into Equation 3.11 for the next iteration. Iterations continue until the desired degree of accuracy, ϵ, is obtained–that is until $|\mathbf{\Delta r}| < \epsilon$, which typically takes less than 5 iterations, even with the suggested initial guesses. Each iteration

leads to a new estimate of the GPS time, t, of signal reception, a new satellite position and transit time, $\mathbf{\Delta p}$, and the geometry matrix $\mathbf{G}$.

The least-squares solution assumes that all of the pseudorange measurements have equal error, which is typically not true in the case of GPS. For example, lower elevation satellites have larger pseudorange errors due to ionospheric and tropospheric propagation errors, and ephemeris errors are typically unequal for the satellites. A weighted least-squares approach would further improve the navigation solution.

$\mathbf{G}$ is typically called the geometry matrix because close inspection shows it to contain the unit vectors to the each of the k satellites:

$$\mathbf{G} = \begin{bmatrix} -\frac{x_s^1 - x_{r_i}}{\rho^1(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{y_s^1 - y_{r_i}}{\rho^1(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{z_s^1 - z_{r_i}}{\rho^1(x_{r_i}, y_{r_i}, z_{r_i})} & -1 \\ -\frac{x_s^2 - x_{r_i}}{\rho^2(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{y_s^2 - y_{r_i}}{\rho^2(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{z_s^2 - z_{r_i}}{\rho^2(x_{r_i}, y_{r_i}, z_{r_i})} & -1 \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{x_s^k - x_{r_i}}{\rho^k(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{y_s^k - y_{r_i}}{\rho^k(x_{r_i}, y_{r_i}, z_{r_i})} & -\frac{z_s^k - z_{r_i}}{\rho^k(x_{r_i}, y_{r_i}, z_{r_i})} & -1 \end{bmatrix} \tag{3.19}$$

The measurement geometry contained in $\mathbf{G}$ allows for estimates of the quality of the RMS positioning error of the navigation solution. The user-satellite geometry has a large impact on the quality of the navigation solution. Consider a simple two-dimensional position estimation made from two satellites, as shown in Figure 3.16. Consider the situation when the angle between the two satellites is obtuse (case A) coupled with a ranging error, ϵ. The area of possible solutions is quite large. A similar situation occurs when the angle between the satellites is obtuse. In case B, where the satellites are at a right angle, the uncertainty in position is minimized. We wish to characterize these uncertainties.

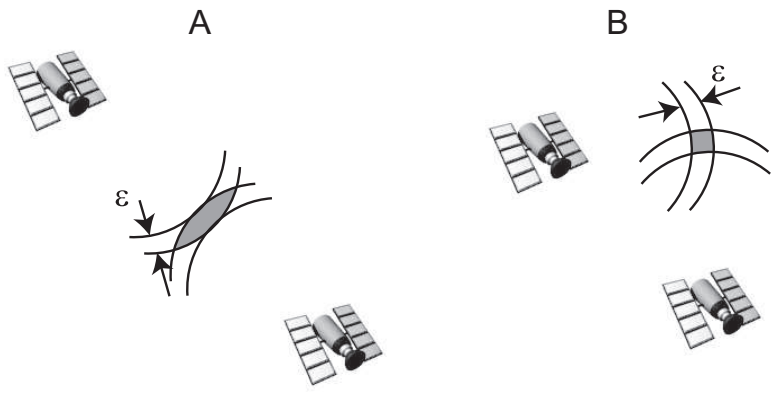

Figure 3.16: Cartoon illustrating a high DOP configuration (a), and a minimal DOP configuration (b).

To relate the quality of the estimates to the geometry matrix, $\mathbf{G}$, it must be assumed that the pseudorange measurement errors are zero-mean gaussian so that the navigation solution error and the receiver clock error are zero-mean gaussian as well. Keeping with the prior definitions, $\mathbf{\Delta r}$ is the error between the best estimate for the navigation solution and the true antenna position and $\Delta c\delta_r$ is the error in the receiver clock offset. The quantities are related to the pseudorange error $\mathbf{\Delta p}$, via Equation 3.18, so that the positional error is given by Equation 3.18. In this case the expected value, $E(\bullet)$, of each quantity is:

$$E(\mathbf{\Delta r}) = 0 \tag{3.20}$$

$$E(\Delta c\delta_r) = 0 \tag{3.21}$$

$$E(\mathbf{\Delta p}) = 0 \tag{3.22}$$

Since the quantities are assumed to be zero-mean gaussian, the pseudorange errors between satellites are uncorrelated, and so the variance of the pseudorange error is:

$$E(\mathbf{\Delta p \Delta p}^\top) = \sigma^2_{\mathbf{\Delta p}} \mathbf{I} \tag{3.23}$$

where $\mathbf{I}$ is the identity matrix. The variance of the navigation solution error can be calculated using Equation 3.18:

$$\begin{aligned} E\left(\begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix}\begin{bmatrix} \mathbf{\Delta r} \\ \Delta c\delta_r \end{bmatrix}^\top\right) &= E\Big(\left(\mathbf{G}^\top\mathbf{G}\right)^{-1}\mathbf{G}^\top\mathbf{\Delta p \Delta p}^\top\mathbf{G}\left(\mathbf{G}^\top\mathbf{G}\right)^{-1}\Big) \\ &= \left(\mathbf{G}^\top\mathbf{G}\right)^{-1}\sigma^2_{\mathbf{\Delta p}} \\ &= \mathbf{Q}\sigma^2_{\mathbf{\Delta p}} \end{aligned} \tag{3.24}$$

where $\mathbf{Q} = \left(\mathbf{G}^\top\mathbf{G}\right)^{-1}$.

The variance in the pseudorange measurements map into navigation solution variance via the user-satellite geometry matrix, $\mathbf{G}$. Since $\sigma^2_{\mathbf{\Delta p}}$ is diagonal, the diagonal elements of $\mathbf{Q}$ scale the variance in the pseudorange measurements to the variance in the navigation solution. $\sigma^2_{\mathbf{\Delta p}}$ is equal to the UERE, introduced in Section 3.3.1. The 3-D position and clock bias estimation error is given by the geometric dilution of precision (GDOP) or:

$$GDOP \cdot \sigma_{\mathbf{\Delta p}} = \sqrt{\mathbf{Q}_{11} + \mathbf{Q}_{22} + \mathbf{Q}_{33} + \mathbf{Q}_{44}}\sigma_{\mathbf{\Delta p}} \tag{3.25}$$

and the positional dilution of precision (the 3-D position estimation error) is given as:

$$PDOP = \sqrt{\mathbf{Q}_{11} + \mathbf{Q}_{22} + \mathbf{Q}_{33}} \tag{3.26}$$

and finally the clock bias estimation error is given by the time dilution of precision as:

$$TDOP = \sqrt{\mathbf{Q}_{44}} \tag{3.27}$$

Other useful dilution of precision (DOP) metrics exist, including Horizontal DOP (HDOP) (related to the East DOP (EDOP) and North DOP (NDOP)), and the Vertical DOP (VDOP). These DOPs are obtained by rotating the $\mathbf{Q}$-matrix into the local vertical-east-north (VEN) reference frame:

$$\tilde{\mathbf{Q}} = \mathbf{R}_{\mathbf{VEN}}\mathbf{Q}\mathbf{R}_{\mathbf{VEN}}^{\top} \tag{3.28}$$

where $\mathbf{R}_{\mathbf{VEN}}$ is the local vertical-east-north rotation matrix. Then:

$$\begin{aligned} EDOP &= \sqrt{\tilde{\mathbf{Q}}_{11}} && (3.29)\\ NDOP &= \sqrt{\tilde{\mathbf{Q}}_{22}} && (3.30)\\ VDOP &= \sqrt{\tilde{\mathbf{Q}}_{33}} && (3.31)\\ HDOP &= \sqrt{\tilde{\mathbf{Q}}_{11} + \tilde{\mathbf{Q}}_{22}} && (3.32)\\ TDOP &= \sqrt{\tilde{\mathbf{Q}}_{44}} && (3.33) \end{aligned}$$

The larger any of the DOP terms, the larger the corresponding navigation errors will be. The pseudorange measurement errors, however, are gaussian distributed and so it is not always true that a large DOP will always result in a large navigation solution error. High DOP, however, will result in larger mean errors. The variance

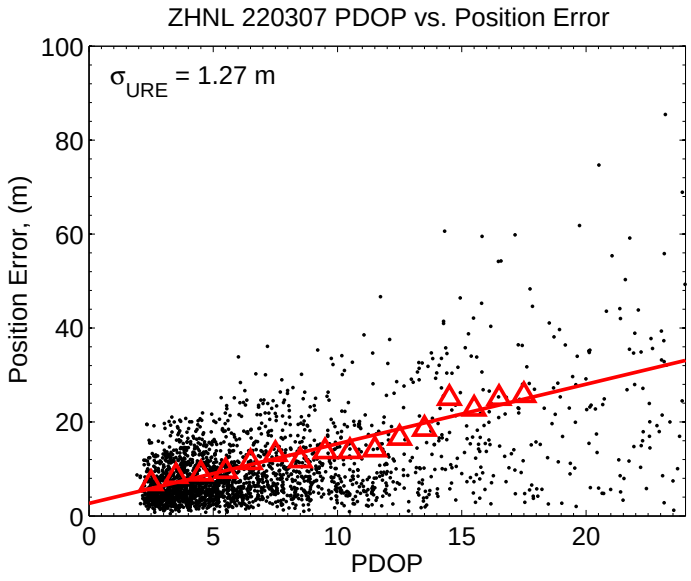

Figure 3.17: Each dot represents a Positional Dilution of Precision (PDOP) versus measured position error. The figure was created from 24 hours of L1 C/A code data from a receiver located in Honolulu, Hawaii at solar minimum (22 March 2007). The sample interval is three minutes, and a random four-satellite selection was chosen to compute the DOP. The red line represents a least-squares fit to binned PDOP and position errors. The slope of the line is approximately σ_{URE}, the user estimated ranging error. The data was compiled using JPL Final Orbital and Clock data (position errors less than 5 cm and clock errors less than 0.1 ns), and JPL IONEX files were used to remove the ionospheric delay (typically 2 - 8 TECU accuracy).

in the position errors is always larger for high DOP than for low DOP. This concept is illustrated in Figure 3.17.

To obtain an intuitive feel for DOP, we turn to Equation 3.24 and $\mathbf{Q} \equiv \left(\mathbf{G}^\top \mathbf{G}\right)^{-1}$. In general, for a square matrix $\mathbf{A}$, its inverse is given as:

$$\mathbf{A}^{-1} = \frac{\text{adj}(\mathbf{A})}{|\mathbf{A}|} \tag{3.34}$$

where adj($\mathbf{A}$) is the adjunct of $\mathbf{A}$. Therefore, to minimize the values of $\mathbf{Q}$, one needs to maximize $|\mathbf{Q}|$. Maximizing the determinant is akin to maximizing the

geometrical volume enclosed by the receiver and all satellites it is tracking. In a four-satellite situation, optimal DOP is achieved by having the antenna at the center of a tetrahedron with the satellites positioned at the four corners of the tetrahedron. This configuration, however, is not possible from any location on Earth. In general, satellites positioned along a great-circle, or grouped together will produce higher values of DOP. Furthermore, choosing satellites satellites at nearly equal elevations will also result in higher DOP. A restricted view of the sky may happen in dense forests or the urban canyons of cities and results in less-than-optimal viewing conditions, higher DOP, and consequently higher positioning errors.

3.8 Conclusion

The chapter began with the introduction of the satellite, control, and user segments, which represent the three major GPS segments. The space segment consists of nominally 24 satellites at 4 R_E orbits that continuously transmit the timing and ranging signals to the users. The control segment receives the signals from the satellites and processes them to predict the satellite orbits and clock drifts and to monitor satellite health. The predictions are uploaded to the satellites every two hours to be re-transmitted as part of the satellite ephemeris in the navigation message. Finally, the user segment receives the satellite signals and processes them within receivers to determine position and time.

Performance requirements set forth by Congress guarantee globally-averaged errors of less than 13 meters horizontally and 22 meters vertically. In recent years, these requirements have been exceeded, and many users have enjoyed accuracies

better than 5 meters. Ranging errors that affect the system include satellite clock and ephemeris errors, ionospheric and tropospheric errors, multipath errors, and other receiver-dependent errors. Today, ionospheric propagation errors are the largest source of error for single frequency users, closely followed by multipath errors.

The GPS signal structure is a code division multiple access system transmitting RHCP signals at L1 (1.57542 GHz) and L2 (1.2276 GHz). The military has encrypted signals that offer greater precision and increased immunity to interference effects. The military user also has access to both the L1 and L2 signals allowing for correction of ionospheric effects. Civilians are primarily restricted to accessing the L1 coarse acquisition sequence, which offers decreased performance. Some civilian semi-codeless receivers are capable of tracking the encrypted L2 signal, but suffer the consequence that their measurements are noisy and prone to loss of lock.

GPS is a weak signal system. The signal strength has implications for its performance in the presence of strong interference. For GPS, the peak received signal power density is at least 10 dB below the noise floor. Once the signal is despread using the local replica of the PRN sequence, the carrier-to-noise ratio for a healthy receiver is typically in the range of 45 - 50 dB-Hz. Most civilian L1 receivers lose lock on the satellite signal when the carrier-to-noise ratios fall below 30 dB-Hz. These metrics are important in future chapters to understand the threshold at which GPS receivers will lose lock in the presence of strong interference signals such as solar radio bursts.

Hardware and software implementations of GPS are explored. These implementations, although drastically different in execution, fundamentally share common design principles and provide similar results. Receiver operation, including the

code and carrier tracking loops are briefly introduced. The impact of interference on the tracking loops will be explored in the next chapter. GPS receiver algorithms, including the navigation solution and the concept of dilution of precision, are introduced to provide necessary background material to understand how the solar radio burst affects positioning accuracy. These concepts will be explored in the coming chapters.

Chapter 4

Solar Radio Burst effects on the Global Positioning System

Whenever you have eliminated the impossible, whatever remains, however improbable, must be the truth. - Arthur Conan Doyle

4.1 Introduction

The ionosphere is considered to be the primary source of space weather effects on Global Positioning System (GPS) signals, systems, and navigation accuracy. This chapter investigates a different class of space weather effects on GPS signals: solar radio bursts. Solar radio bursts affect the GPS receiver by lowering the received carrier-to-noise ratio.

As discussed in Chapter 2, solar radio bursts are intense radio emissions from the sun that are usually associated with solar flares (*Kundu*, 1965; *Bastian et al.*, 1998). Typically, incoherent gyrosynchrotron emissions dominate as the cause of these radio emissions at centimeter and millimeter wavelengths, whereas coherent plasma radiation is often the cause at meter and decimeter wavelengths.

Plasma radiation occurs when electron beam energy is nonlinearly converted to plasma waves at harmonics of the local plasma frequency, which are then converted to ordinary-mode electromagnetic waves. Typically, these radio bursts occur in the frequency range from a few tens of megahertz up to 3 GHz (*Bastian et al.*, 1998).

*Portions of this chapter are based on the publication *Cerruti et al.* (2006), which is joint collaborative work with A. P. Cerruti, P. M. Kintner, D. E. Gary, L. J. Lanzerotti, E. R. de Paula, and H. B. Vo. Reproduced by permission of the American Geophysical Union.

Plasma radiation is the primary mechanism that causes solar radio bursts at the GPS frequencies.

The polarization of any burst is determined by the local magnetic field at the mode conversion site. Since there is no preferred direction to the magnetic field, the solar radio burst can be right-hand circularly polarized, left-hand circularly polarized, or a mixture of both when viewed with respect to its wave vector at earth.

Solar radio bursts can have durations from tens of seconds to a few hours. Typically, solar radio bursts that are concurrent with a solar flare are impulsive, while those that occur in the decay phase are smoother and of longer duration. Over the duration of the burst, the peak frequency can climb, fall, or both. Furthermore, the intensity and polarization of the burst can vary. Finally, the bandwidth of the solar radio burst at any point in time can vary substantially.

Solar radio bursts affect the system by increasing the noise floor of the receiver, which in turn causes a decrease in C/N_0. In this chapter, an explanation and computational analysis will be given to detail how a solar radio burst affects the receiver. Analysis of a moderate solar radio burst event in the GPS frequencies from September 2005 will be given, as well as an analysis of the record-setting events that occurred in December 2006. The events show that the Global Positioning System can be drastically impacted by solar radio bursts.

4.2 Solar Radio Burst Induced Degradation of Carrier-to-Noise Ratio

Since GPS is a right-hand circularly-polarized system, most GPS antennas are built to be sensitive only to right-hand circularly-polarized radiation (see Sections 3.4 and 3.6). Most GPS receivers contain an automatic gain control (AGC) to keep the incoming analog signal within the proper voltage levels for the analog-to-digital converter. Since the peak GPS signal power lies below the noise (Section 3.6), the AGC will drive the gain such that the quantized noise stays within predefined limits. In-band RHCP solar radio bursts are an unwanted signal that causes an increase in the spectral noise density of a GPS receiver, which drives down the gain of the AGC, thus decreasing the value of C/N_0. As long as the total incoming signal does not drive the AGC beyond its limits, the AGC should remain within a linear mode of operation. Furthermore, the AGC response time is typically fast (a few milliseconds) and will adjust to the quickly changing solar radio burst power. Over the duration of the solar radio burst, the satellite itself is not affected, and in fact, the transmitted signal power remains unchanged.

Solar radio bursts are not the only source of signal degradation of GPS carrier-to-noise ratio. Significant effort has been devoted to investigating the effects of ionospheric scintillations on the GPS signal. Ionospheric scintillations cause large swings in the received signal power, C that are concurrent with rapid changes in signal phase. Amplitude scintillation are a diffractive process in which the phase of the incoming GPS signal is modified by irregularities in the ionosphere resulting in an interference pattern on the ground. The interference pattern may have a variation in the carrier-to-noise ratio of more than 20 dB in some cases with both increases and decreases of C/N_0 with respect to the nominal, expected C/N_0. The

variations in C/N_0 can lead to loss of lock on the scintillating signals. At times, the phase of the incoming signal varies rapidly, which also contributes to cycle slips and loss of carrier lock.

Solar radio bursts have a similar effect on the GPS receiver operation as ionospheric scintillation. Solar radio bursts, however, are generally a wide-band interfering radio source from the sun that causes increases in the spectral noise density, N_0, of the receiver. The carrier power, C, from the satellite remains constant during the solar radio burst. As during scintillation, the carrier-to-noise ratio during a solar radio burst decreases, although with different time scales, and never with an associated increase in the carrier-to-noise ratio. In addition there is no reason to expect rapid phase fluctuations during the solar radio burst.

4.2.1 L1 Degradation

Following the *Klobuchar et al.* (1999) analysis, a theoretical value for solar radio burst power that will cause a 3 dB drop in the carrier-to-noise ratio at L1 can be calculated. The following analysis is similar to the calculation for the noise floor calculated in Section 3.6. First the equivalent noise temperature of the receiver and antenna must be determined. Depending on the antenna pattern and the amount of ground temperature observed, a typical GPS antenna has a source (sky) noise temperature of 75 - 100 K (*Van Dierendonck*, 1996). The total system noise temperature is computed as:

$$T_{system} = T_{source} + T_{receiver} \text{ K} \tag{4.1}$$

A typical noise figure, NF, for a GPS receiver is 4 dB (2.51). For receivers with a high-gain low-noise amplifier located in the antenna, the noise figure is set by noise figure of the low-noise amplifier. For receivers with a high first stage gain and negligible line losses, the equivalent receiver noise temperature is approximately:

$$\begin{aligned} T_{receiver} &= T_{ambient} \times (NF - 1) \\ &= 290 \text{ K} \times (2.51 - 1) \\ &= 438 \text{ K} \end{aligned} \tag{4.2}$$

Where $T_{ambient} = 290K$ is the ambient temperature of the receiver. From Equation 4.1, the equivalent system temperature is:

$$\begin{aligned} T_{system} &= 100 \text{ K} + 438 \text{ K} \\ &= 538 \text{ K} \end{aligned} \tag{4.3}$$

The power density, P_n, of the system temperature is determined by multiplying T_{system} by Boltzmann's constant ($k_B = 1.3806503 \cdot 10^{-23} \ \frac{\text{m}^2\text{kg}}{\text{s}^2\text{K}}$):

$$\begin{aligned} P_n &= T_{system} \times k_B \\ &= 7.42 \cdot 10^{-21} \text{ W/Hz} \end{aligned} \tag{4.4}$$

The signal power collected by a GPS antenna depends on the effective area of the antenna. The effective area of an antenna is related to the gain, g, and the operating wavelength, λ ($\lambda = 0.1903$ m at GPS L1). The gain of most hemispherical

GPS antennas is 3 - 5 dBic (isotropic circular) or higher at zenith and decreases as the elevation decreases. A gain of 4 dBic (2.51) will be used for the analysis, which differs from that of *Klobuchar et al.* (1999) because they assumed an antenna gain of 1 dBic. The effective area of the antenna is calculated as:

$$\begin{aligned} A_{eff} &= \frac{g\lambda^2}{4\pi} \\ &= \frac{2.51 \times 0.19029^2}{4\pi} \\ &= 7.23 \cdot 10^{-3} \text{ m}^2 \end{aligned} \tag{4.5}$$

To cause a 3 dB drop in the carrier-to-noise ratio, the incident RHCP power received from the solar radio burst must equal the noise floor power of the receiver. Therefore, dividing the system noise temperature, T_{system}, by the antenna gain, A_{eff}, yields the equivalent solar radio burst power, P_{SRB}, required to cause a 3 dB loss in the carrier-to-noise ratio:

$$\begin{aligned} P_{SRB} &= \frac{P_n}{A_{eff}} \\ &= \frac{7.42 \cdot 10^{-21} \text{ W/Hz}}{7.23 \cdot 10^{-3} \text{ m}^2} \\ &= 1.03 \cdot 10^{-18} \text{ W/m}^2\text{−Hz} \end{aligned} \tag{4.6}$$

Finally, one solar flux unit (SFU) is equal to 10^{-22} W/m^2-Hz, so the equivalent solar radio burst power, P_{SFU}, in solar flux units is:

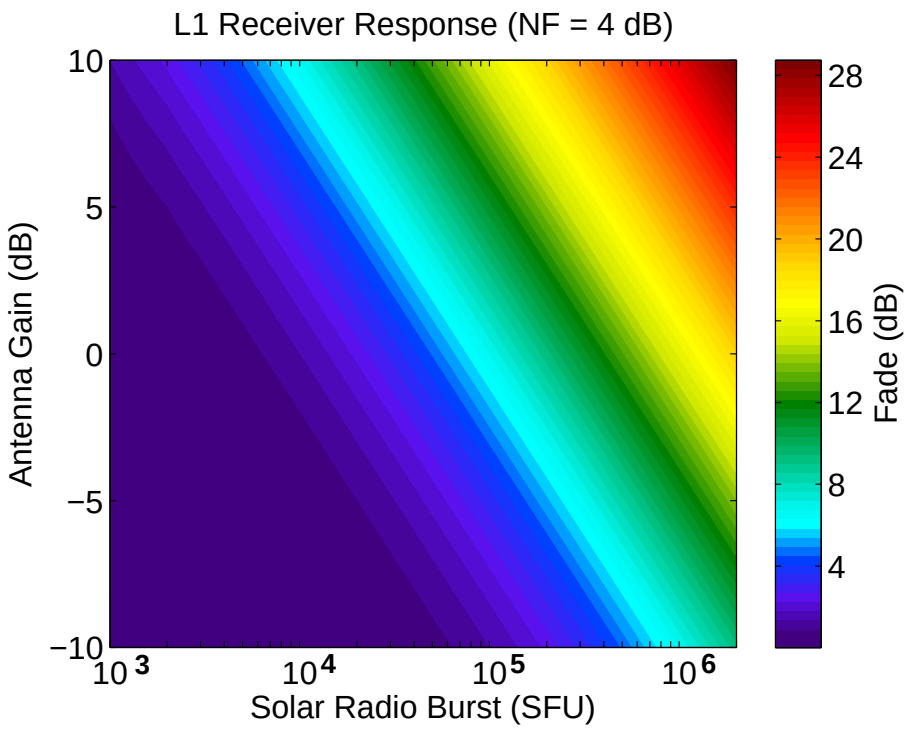

Figure 4.1: A plot demonstrating the expected fade at L1 given a solar radio burst intensity and antenna gain. The plot is for a receiver operating at 25°C with a noise figure of 4 dB.

$$
\begin{aligned}
P_{SFU} &= \frac{P_{SRB}}{10^{-22}\dfrac{\mathrm{W/m^2{-}Hz}}{SFU}} \\
&= \frac{1.03 \cdot 10^{-18}\ \mathrm{W/m^2{-}Hz}}{10^{-22}\dfrac{\mathrm{W/m^2{-}Hz}}{SFU}} \\
&= 10,300\ \mathrm{SFU}
\end{aligned}
\tag{4.7}
$$

Thus, at L1, with an antenna gain of 4 dBic, a solar radio burst of 10,300 SFU RHCP will cause a 3 dB C/N_0 fade. Figure 4.1 illustrates the magnitude of a fade expected for differing antenna gains at a given solar radio burst intensity.

4.2.2 L2 degradation

An analysis similar to the one conducted for GPS L1 was also made to determine the solar radio burst power that would cause a 3 dB fade on GPS L2 (1.2276 GHz). Assuming an antenna gain of 4 dBic at L2, a 3 dB C/N_0 fade will occur for a 6,218 SFU RHCP solar radio burst. For equal antenna gain, the receiver is more sensitive to the solar radio burst merely because the effective collecting area of the antenna is larger at longer wavelengths. The P(Y) signal transmitted at L2 is 3 dB lower in power than the P(Y) signal at L1, rendering the L2 signal more vulnerable to solar radio bursts. Figure 4.2 shows the response of the receiver to a solar radio burst as a function of antenna gain.

Typically dual-band (L1 and L2) GPS antennas have significantly lower gain at L2 than they do at L1. For example, the NovAtel GPS-702L has a peak gain at zenith of 5 dBic on L1 but only a peak gain of 1.5 dBic at L2. The difference in gain is not troublesome for GPS receivers because the longer-sequence military codes do not require as much power as the L1 C/A code to be adequately tracked. The difference in this particular antenna's gain clearly implies that the L2 channel will observe a smaller fade at L2 than at L1 for a solar radio burst of equal intensity at both frequencies.

Most users do not have access to the encrypted military codes on the L2 signal. Instead, codeless, or semi-codeless receivers are utilized. Most receivers rely on some variant of a cross-correlation technique, which correlates the L1 P(Y) signal against the L2 P(Y) signal, effectively raising the noise floor. The noise floor on the L2 portion of a cross-correlating semi-codeless receiver will be the sum (in dB) of the noise floor from the L1 and the L2 signal paths. Furthermore, L2 semi-codeless receivers are particularly susceptible to loss of carrier tracking and

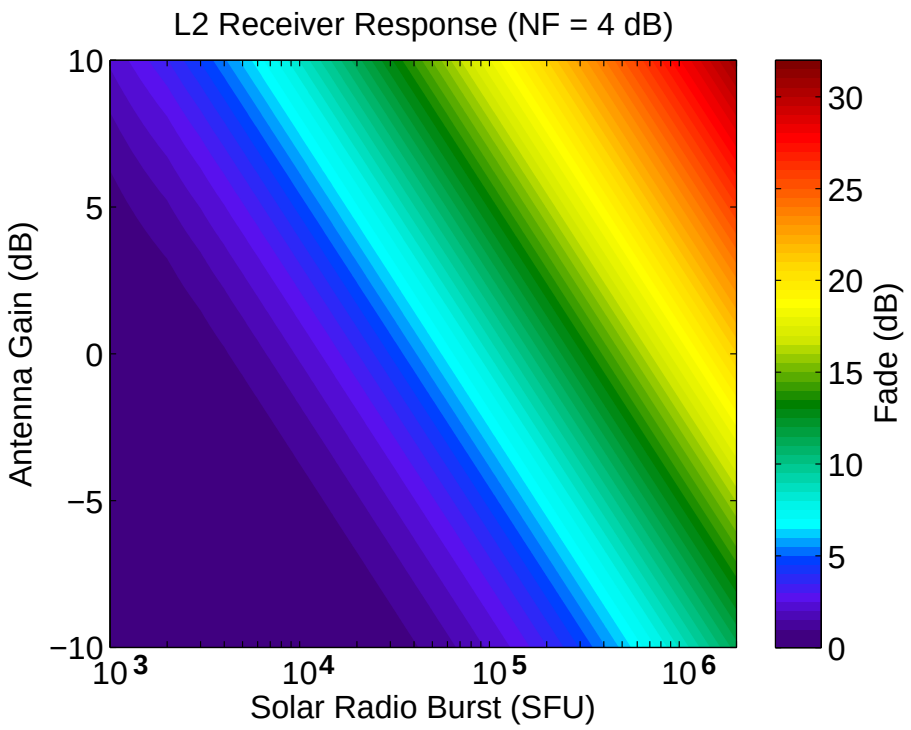

Figure 4.2: A plot demonstrating the expected fade at L2 (using a coded receiver) given a solar radio burst intensity and antenna gain. The plot is for a receiver operating at 25°C with a noise figure of 4 dB. The color scale is the same as in Figure 4.1.

cycle slips during signal fades because of squaring loss that occurs in the receiver tracking loops. Squaring loss is the decrease in signal-to-noise ratio within the tracking loops of the receiver because the receiver must square the incoming signal to remove the data-bits from the signal. As a result of the squaring action, the noise in the in-phase and quadrature channels of the loops are multiplied together, effectively squaring the noise.

The decreased performance of semi-codeless and code-less receivers during solar radio bursts is analyzed by *Chen et al.* (2005). In the paper, they show a correlation coefficient of 0.75 between loss of carrier lock on the L2 signal of semi-codeless receivers and solar radio bursts recorded at 1.415 GHz for the 28 October 2003 solar radio burst associated with the X17.2 solar flare at 1102 UT. From the same

event, they determine that the threshold for which semi-codeless receivers suffer from loss of lock occurs at roughly 4,000 - 12,000 SFU (total polarization). *Chen et al.* (2005) also show a clear correlation between increased occurrence of loss of lock for receivers located near the subsolar point, but do not provide an explanation for this observation.

Guided by the original analysis of *Klobuchar et al.* (1999), *Chen et al.* (2005) determined a loss of carrier lock threshold of about 20,000 - 100,000 SFU for a semi-codeless receiver when operating with an antenna gain of -3 dBic, depending on the incident carrier-to-noise ratio. Although the authors show a clear correlation between loss of carrier lock and solar radio bursts in the GPS frequencies, they do not show the effect the solar radio burst had on the received carrier-to-noise ratio, nor do they adequately explain why their estimates are in disagreement with their observations. The authors simply conclude that the threshold at which solar radio bursts impact GPS receivers must be re-evaluated, as compared to the threshold determined by *Klobuchar et al.* (1999), which was determined for L1 C/A code receivers. Loss of lock on a receiver is a function of the tracking loop parameters implemented within each receiver and are explored in the next section.

4.2.3 Receiver Tracking Errors

The motivation for this section is to determine at what point a GPS receiver will lose lock as a result of decreased carrier-to-noise ratio. Inside every GPS receiver are two tracking loops operating in tandem that are responsible for maintaining signal lock. The first is a delay-lock loop (DLL), which maintains the phase alignment of the locally-generated PRN sequence with that arriving from the satellite. The second is a phase lock loop (PLL) that aligns the phase of the locally gener-

ated carrier wave with that arriving from the satellite. Failure of either tracking loop will cause the receiver to lose code or carrier lock on that satellite and will result in degraded performance. Fortunately, each can be analyzed independently, and the PLL performance is explored first.

Since GPS utilizes a system whereby the carrier is modulated by a Pseudo-random noise (PRN) sequence and by navigation data bits, a non-coherent discriminator must be used in the dual Costas phase lock loop (PLL). This type of PLL allows the loop to ignore the phase inversions induced by the PRN sequence and navigation data bits and ultimately recover the carrier phase. A discriminator is merely a non-linear mathematical function that has been carefully chosen to expose the parameter to be estimated (e.g. phase for the PLL or code delay for the DLL) while suppressing the effects of other parameters (e.g. noise, or bit transitions) (*Misra and Enge*, 2006). A discriminator commonly used in GPS for the PLL is simply $\tan^{-1}(Q/I)$, where I and Q are the In-phase and Quadrature power estimates of the signal. Other discriminators are discussed in *Van Dierendonck* (1996) and *Ward* (1996a).

The process of recovering the carrier from the GPS signal suffers from a degradation of signal-to-noise in the loop bandwidth of the PLL. This degradation, known as squaring loss, challenges the ability of the receiver to track a signal in weak-signal conditions, as compared to an ideal phase-lock loop, which only tracks a continuous sine-wave signal[1]. A linear model for the 1-σ phase error of the Costas loop, σ_ϕ, is given by:

[1]For more details on Costas loop tracking, please refer to *Simon and Lindsey* (1977) and *Simon* (1978).

$$\sigma_\phi = \frac{180}{\pi}\sqrt{\frac{B_L}{C/N_0}S_L} \text{ degrees} \tag{4.8}$$

and is related to the squaring loss factor, S_L, the carrier-to-noise ratio C/N_0 of the incoming signal, and the single-sided noise loop bandwidth, B_L (Hz) (*Woo*, 2000). For the $\tan^{-1}(Q/I)$ discriminator, the linear model is valid for angles up to about 45°. However, only phase errors less than 15° are tolerable (*Ward*, 1996a).

For a continuous-wave (no data-bit modulation) signal, when a linear Costas loop is utilized, S_L is unity. Otherwise, S_L is calculated as:

$$S_L \cong 1 + \frac{B_P}{2C/N_0} \tag{4.9}$$

where B_P represents the pre-correlation bandwidth of the signal, and C/N_0 is again the carrier-to-noise ratio.

For the known, L1 C/A code receiver, the effective bandwidth of the C/A code, which is approximately 2 MHz, is reduced by integrating over multiples of the C/A code period (1 ms) up to the bit period (20 ms). Therefore, B_P, varies from 50 - 1000 Hz. On the other hand, the L2 P(Y) signal is essentially unknown to the civilian user, so B_P is approximately 20 MHz, the two-sided bandwidth of the 10.23 MHz wide P(Y)-code.

Figure 4.3 illustrates the tracking error (top panel) and the squaring loss terms (bottom panel) for various tracking loops as a function of carrier-to-noise ratio. The black-dashed line represents the ideal Costas loop with $B_L = 15$ Hz, which suffers no squaring loss. The gray line represents an L1 C/A code with $B_L = 15$ Hz and $B_P = 1000$ Hz (1 ms integration). Finally, the black solid line represents

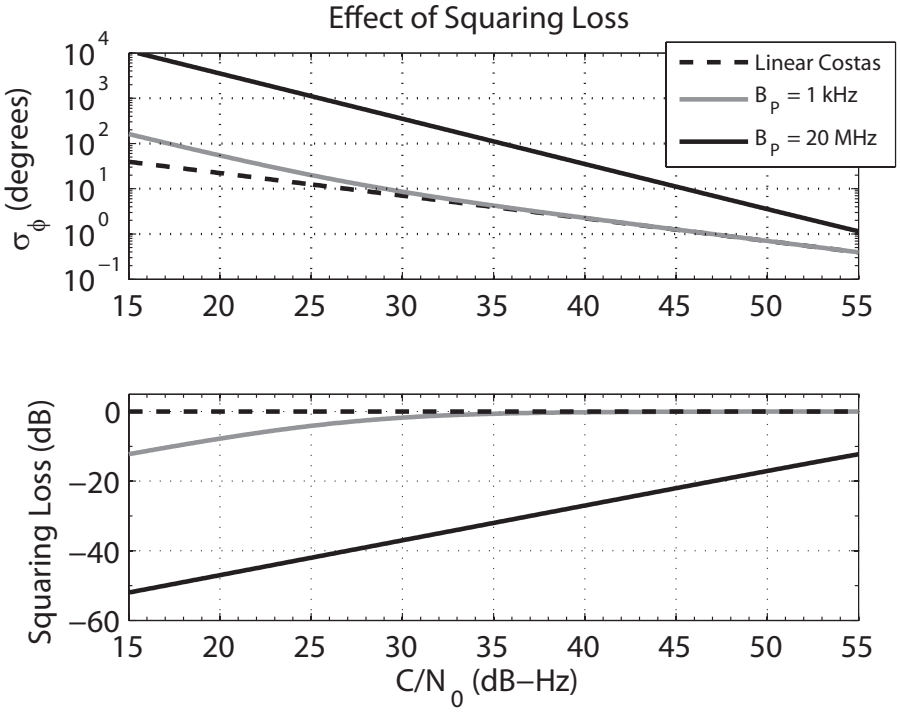

Figure 4.3: The top panel illustrates the RMS tracking error in degrees as a function of carrier-to-noise ratio for various Costas loops, assuming a loop bandwidth, B_L, of 15 Hz. The bottom panel illustrates the squaring loss, which is independent of tracking loop bandwidth.

the L2 P(Y) semi-codeless technique with $B_L = 15$ Hz and $B_P = 20$ MHz. The tracking error for the standard L1 C/A code Costas loop diverges substantially from the linear Costas loop (black dashed line) at about 30 dB-Hz. According to *Ward* (1996a), the maximum tolerable 1-σ_ϕ phase error before cycle slips occurs is 15°. For the L1 C/A code, with $B_P = 1000$ Hz, this threshold occurs at 26.6 dB-Hz. At $B_P = 50$ Hz, the threshold decreases to 23.8 dB-Hz. It should be noted that the loop bandwidth, B_L, could be decreased substantially thus further decreasing the loss of lock threshold. Decreasing B_L comes with a cost: the receiver will suffer from degraded dynamic stress performance (i.e. receiver-satellite relative acceleration or jerk and ionospheric induced acceleration or jerk), which can also cause a loss of lock.

The maximum tolerable phase error of 15° occurs at an L2 carrier-to-noise ratio of approximately 43.7 dB-Hz, which corresponds to an L1 C/A code carrier-to-noise ratio of 46.7 dB-Hz. The tracking error of the L2 P(Y) semi-codeless technique is substantially higher than the other techniques because the bandwidth, B_P, of the L2 P(Y) semi-codeless technique is much larger than the L1 C/A code. Squaring loss is significant at all carrier-to-noise ratios, as illustrated in the bottom panel. Other semi-codeless techniques for tracking the P(Y) signal exist and improve squaring loss by 10 - 15 dB. The interested reader should consult *Woo* (2000).

As a parting comment on PLL tracking error, it should be noted that the results explored here only take into account phase error due to thermal noise, and not the phase error due to oscillator phase noise (a function of oscillator temperature and quality), dynamic stress (a function of satellite-receiver relative motion and ionospheric dynamics), and oscillator acceleration stress (caused by oscillator drifts). In reality these stresses would factor into the RMS tracking error and should be accounted for. These concepts will be explored in Chapter 6.

The Delay Lock Loop (DLL) maintains the phase of the locally generated PRN sequence to the sequence arriving from the satellite. The DLL ultimately allows ranging measurements to the satellite to be made and is a critical component of a GPS receiver. Furthermore, and perhaps more critically, proper code phase alignment allows the PRN sequence to be stripped from the incoming signal, allowing the PLL to maintain carrier tracking. For the analysis, only a Delay Lock Loop utilizing dedicated early and late correlators and discriminators will be considered. Dot-product discriminators and tau-dithered early/late correlators will not be considered and are discussed by *Ward* (1996a).

The DLL assimilates measurements from three separate correlators: a prompt

correlator, which is a measure of the In-phase and Quadrature (I and Q) signals with the locally generated PRN sequence aligned to the sequence arriving from the satellite, an early correlator, which is a measure of the the I and Q signals with the locally generated PRN sequence shifted ahead of the prompt correlator by a fraction of a PRN code period $d/2$, and finally the late correlator, which is a measure of the I and Q signals with locally generated PRN sequence shifted behind the prompt correlator by a fraction of a PRN code period, $d/2$. In other words, the early and late correlators are shifted the same amount from the prompt correlator, resulting in a total delay between the early and late correlators of d code periods (or chips).

The magnitude of the early, late, and prompt correlators allows the DLL to ascertain the alignment error of the PRN sequence. For example, if the early and late correlations are of equal magnitude and the prompt correlation is of magnitude indicating that the signal is tracking, the locally generated sequence is aligned with the satellite sequence. If the early correlation is less than the late correlation, then the locally generated sequence is early with respect to the satellite sequence. The converse is true if the late correlation is less than the early correlation. Once the locally generated code is misaligned by more than d code periods, the DLL loses lock.

For the early/late type discriminator and correlator, the DLL thermal noise code tracking error (in meters) is:

$$\sigma_t = \lambda_c \sqrt{\frac{2d^2 B_N}{C/N_0} \left[2(1-d) + \frac{4d}{TC/N_0} \right]} \text{ m} \tag{4.10}$$

where λ_c is the code period wavelength in meters, d is the early/late correlator

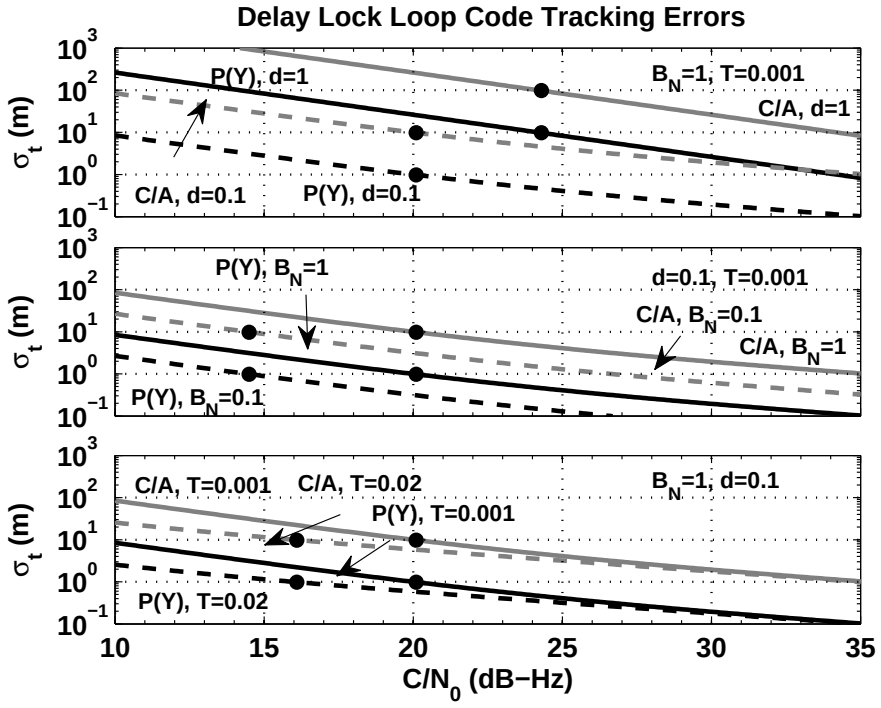

Figure 4.4: Each panel illustrates the RMS code tracking error in meters as a function of carrier-to-noise ratio for various Delay Lock Loop tracking parameters. The point at which the code tracking error exceeds the lock threshold for each signal is denoted by the black dots on each curve.

spacing in code periods, B_N is the DLL code loop noise bandwidth (Hz), C/N_0 is the carrier-to-noise ratio (Hz), and $T = 1/B_P$ is the predetection integration time in seconds (*Ward*, 1996a). The term in brackets corresponds to the squaring loss, which also affects the DLL. According to (*Ward*, 1996a), the 3-σ_t tracking error cannot exceed d code periods, so the 1-σ_t tracking error cannot exceed $d/3$ code periods.

Figure 4.4 illustrates the code tracking errors for various schemes of a rate-aided DLL. A rate-aided DLL merely incorporates estimates of doppler shift from the PLL and allows for substantial reduction of B_N. λ_c for the C/A code is 293.05 m and λ_c for the P(Y) code is 29.305 m. In the top panel, the code period spacing, d is varied, holding $B_N = 1$ Hz and $T = 0.001$ s constant. Typical code period

spacings can range from 0.025 - 1 code periods. Smaller code period spacings reduce multipath (especially on the L1 C/A code), but at the expense of reduced dynamic stress threshold. For most applications, however, rate-aided DLLs allow dynamic stress to be neglected.

The middle panel illustrates the tracking error for various noise bandwidths, B_N, with $d = 0.1$ s and $T = 0.001$ s. In a rate-aided DLL, the code loop noise bandwidth can range from 0.005 - 1 Hz and is determined by dynamics and desired DLL performance. Finally, the bottom panel illustrates the code tracking error when T is varied, holding $B_N = 1$ Hz and $d = 0.1$ s. T cannot exceed the data bit period of the sequence (20 ms for GPS). The carrier-to-noise ratio where loss of code lock occurs is independent of the PRN sequence because the loss of lock point occurs at $d/3$ code periods.

For reasonable implementations of the rate-aided DLL, the code tracking error exceeds the lock threshold at 15 - 25 dB-Hz for both the P(Y) and C/A code sequences. The quality of the code-range navigation solution (i.e. the solution determined from pseudoranges) will deteriorate rather quickly for low carrier-to-noise ratios because σ_t climbs rapidly. At these carrier-to-noise ratios, however, the PLL has usually lost carrier lock. The PLL is the weakest link of the GPS receiver.

The performance of GPS receivers in regards to carrier and code tracking error will be explored in depth in Section 5.3. First however, the computational analysis of Section 4.2 will be verified using a moderate solar radio burst that occurred on 7 September 2005.

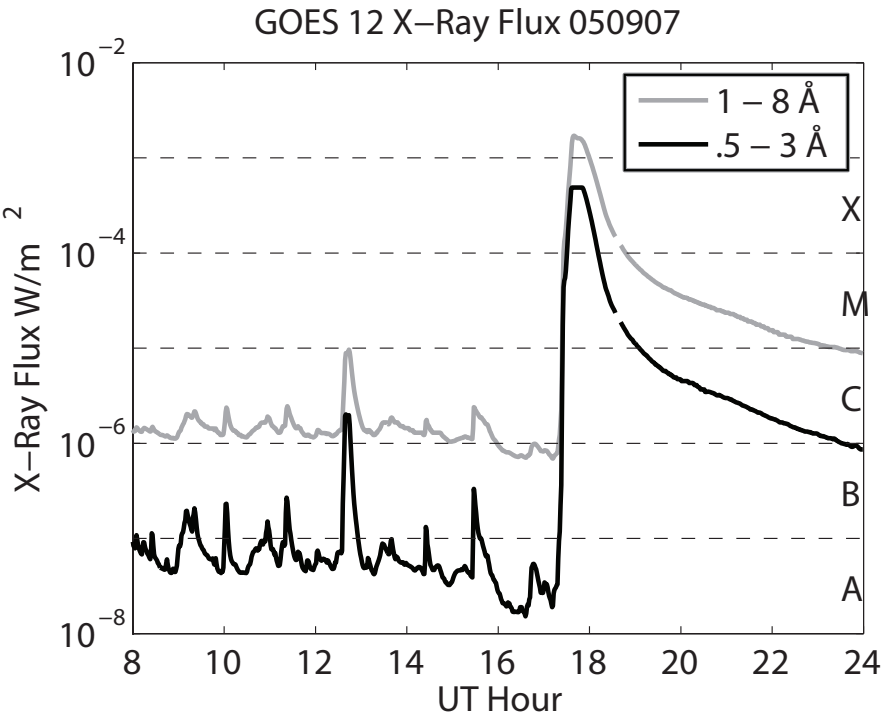

Figure 4.5: X-ray flux from the GOES-12 satellite for 7 September 2005 showing the X-17.2 flare–the fourth largest in recorded history. The flare had a very sudden onset at about 17.6 UT and a slow decay afterward.

4.3 7 September 2005 Solar Radio Bursts

On 7 September 2005 at approximately 1736 UT, Active Region 10798 unleashed a soft x-ray class X17.2 solar flare that was the fourth largest recorded by the GOES satellite constellation. The 0.5 - 3 Å (hard) and 1 - 8 Å (soft) solar x-ray flux is shown in Figure 4.5. Although the flare was exceptionally energetic in the x-ray wavelengths, the associated solar radio burst was quite modest.

Solar radio burst data for this event are acquired using the Owens Valley Solar Array (OVSA), operated by the New Jersey Institute of Technology (located at 37.2°N, 117.3°W). The data are acquired using two 27 meter antennas with log-periodic feeds that record solar radio burst power from 1.2 - 18 GHz. Data are recorded at an 8.1 s sampling rate over a double side-band of 50 MHz bandwidth

each center frequency. For example, the 1.2 GHz measurements are from double side-bands at 1.143 - 1.193 GHz and 1.207 - 1.257 GHz. The OVSA data provide power measurements in the left-hand circular, right-hand circular, and linear polarization modes. Each day the instrument is calibrated, and the measurements are accurate to within 20% (*Nita et al.*, 2004).

Solar Radio Burst data obtained from the OVSA are shown in Figures 4.6 through 4.8. Figure 4.6 shows the total solar radio burst power spectral density from 1.2 - 18 GHz between 1715 and 1848 UT, whereas Figures 4.7 and 4.8 show the solar radio burst power spectral density for the left-hand circular polarized (LHCP) and right-hand circular polarized (RHCP) modes of the spectra, respectively. The power scales in Figures 4.7 and 4.8 are identical.

The initial burst at 1736 UT was a series of three impulses of duration from 1 - 3 minutes, while the secondary burst at 1830 UT was about 10 minutes in duration and exhibited a fairly constant power spectral density. In general, the power in the RHCP and LHCP modes is not the same. For example, the power of the solar radio burst in the LHCP mode at 1830 UT is about twice the power in the RHCP mode. However, for the event at 1736 UT, the power is roughly equal in both modes.

4.3.1 L1 Receiver Response

Concurrent with the solar radio burst data are GPS L1 signal amplitude data recorded using three Cornell ScintMon receivers (*Beach and Kintner*, 2001) located at the Arecibo Observatory in Puerto Rico (18.3°N, 66.8°W). These receivers record signal amplitude data at 50 samples per second from both Global Position-

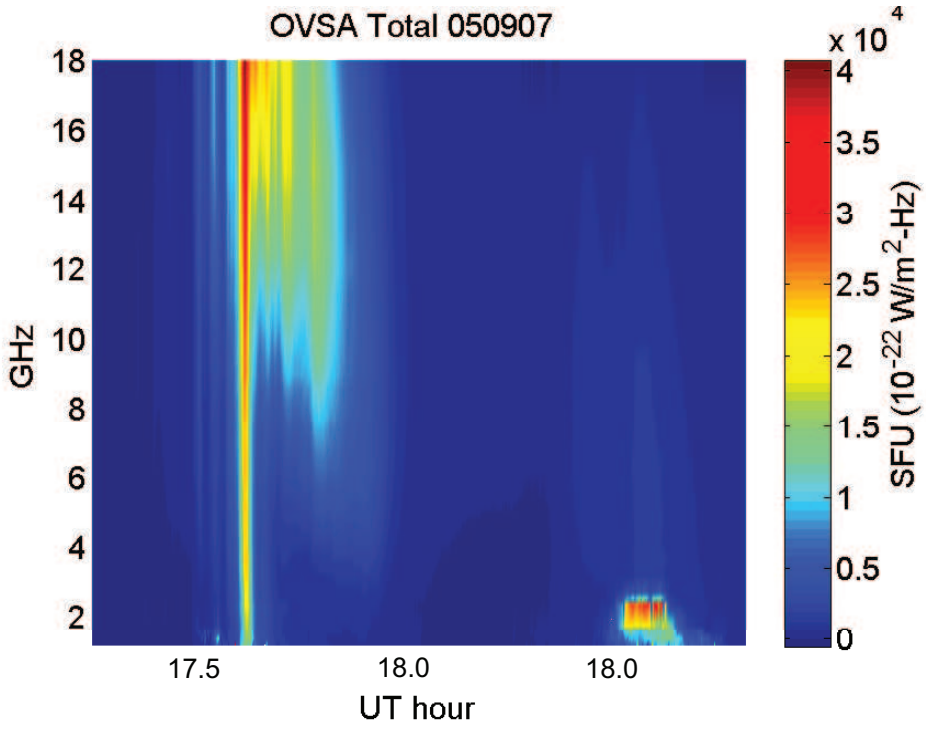

Figure 4.6: Total solar radio burst power spectral density from 1.2 - 18 GHz on 7 September 2005 as measured by the OVSA. There are two periods of activity at 1736 UT and at 1830 UT.

ing System (GPS) and Wide Area Augmentation System (WAAS) satellites. The receivers acquire the signals using an Impact Power GPS-MAM35 RHCP hemispherical antenna mounted on an 18-inch square ground plane. The peak gain of the antenna zenith is 4 dBic, but inclusion of the large ground plane likely increases the peak gain. The GPS signals are processed using Zarlink GP2015 and GP2021 chipsets in a Cornell ScintMon receiver (*Beach and Kintner*, 2001).

Plotted as a black line in the top panel of Figure 4.9 is the carrier-to-noise ratio (C/N_0) of PRN 25 from 1715 - 1848 UT for the quiet day prior to the solar radio burst (6 September 2005). Also plotted as a gray line for the same time period and satellite is the C/N_0 for the day of the burst (7 September 2005), when the satellite elevation varied from 43.4° to 44.2°. The C/N_0 shown is the average over 1 second intervals. The power spectral density of the solar radio bursts in

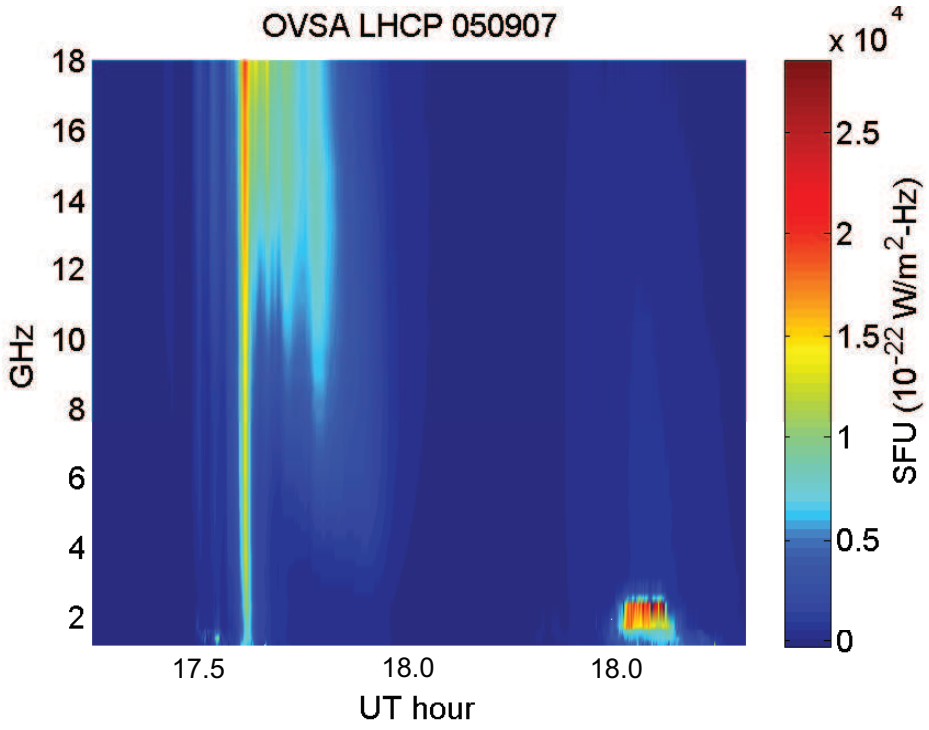

Figure 4.7: Solar radio burst power spectral density in the left-hand circularly polarized mode. Note that the burst at 1830 UT had most of its power in the LHCP mode as opposed to the RHCP mode, shown in Figure 4.8.

the total polarization (black dashed), LHCP (gray), and RHCP (black) modes as measured at 1,600 MHz using the OVSA are shown in the bottom panel of Figure 4.9. Clearly evident in the PRN 25 data for 7 September 2005 at 1736 UT and again at 1830 UT are the decreases in the measured carrier-to-noise ratio due to the solar radio bursts.

For the solar radio burst event at 1736 UT, the total peak power spectral density is 17,000 solar flux units (1 SFU = 10^{-22} W/m^2-Hz). In the LHCP mode, the peak power spectral density was 10,900 SFU, and in the RHCP mode the peak power spectral density was 8,700 SFU. For the event at 1830 UT, the peak total power spectral density was 14,600 SFU, and the LHCP and RHCP peak power spectral densities were 10,900 SFU and 5,400 SFU, respectively.

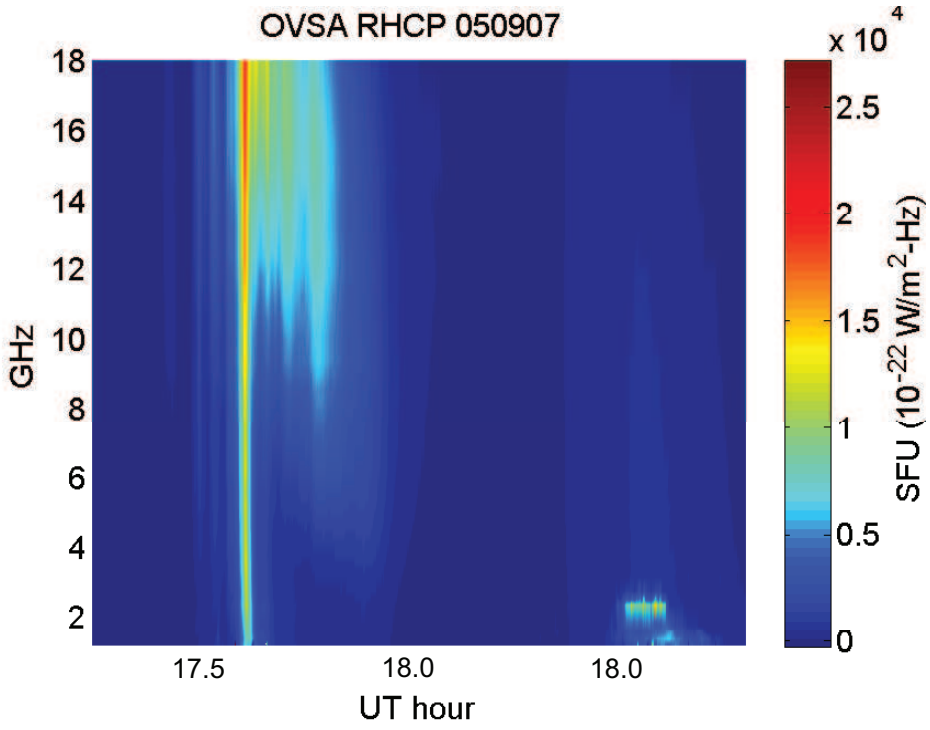

Figure 4.8: Solar radio burst power spectral density in the right-hand circularly polarized mode. The initial burst at 1736 UT had nearly equal power in the RHCP mode as in the LHCP mode. RHCP is the polarization mode that will affect Global Navigation Satellite Systems because of their design.

GPS orbits have a period of 12 sidereal hours, so they occupy the same position in the sky relative to a stationary receiver every sidereal day. For this reason, a similar multipath environment will be repeated from day to day. In Figure 4.9 the fluctuations in carrier-to-noise ratio due to multipath are clearly evident. For example, the large modulations observed in the signal powers on both days at 1815 UT are caused by multipath. By simply subtracting the carrier-to-noise ratio measurements of the quiet previous day shifted by a sidereal day from the carrier-to-noise measurements for the burst day, both multipath variations and expected variations due to satellite-receiver geometry can be removed from the C/N_0 record, yielding a measure of fading, $\Delta C/N_0$. A more positive value of $\Delta C/N_0$ indicates a larger fade (a smaller received carrier-to-noise ratio). Henceforth, any reference

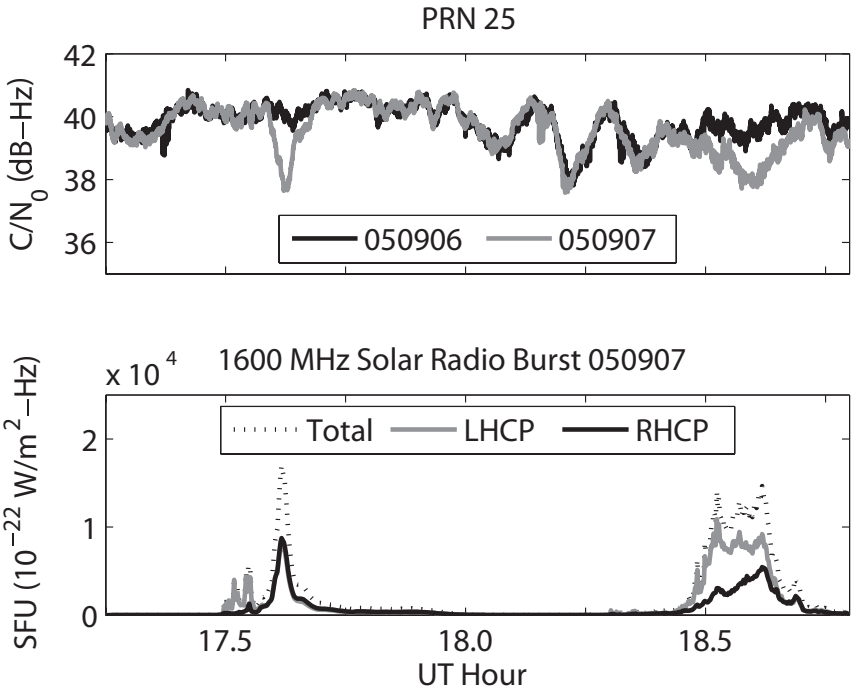

Figure 4.9: The top panel shows the C/N_0 from PRN 25 for the quiet day before the solar radio burst (black), and the day of the burst (gray) as measured using receivers at Arecibo. The bottom panel shows the total, LHCP, and RHCP solar radio burst power spectral density at 1,600 MHz as measured at the OVSA. Notice the fades in signal power that coincide with the solar radio burst activity. The fluctuations in signal power at 1815 UT are caused by multipath and repeats with a period of a sidereal day.

to a quantity $\Delta C/N_0$ or a C/N_0 fade is determined using this method.

In Figure 4.10, the C/N_0 fade ($\Delta C/N_0$) as calculated using this method is plotted (gray) against the total polarization, LHCP, and RHCP solar radio burst power spectral densities at 1,600 MHz (black). All scales are equal. The peak C/N_0 fade at 1736 UT was 2.3 dB, when the sun was at an elevation of approximately 68° with respect to the antenna.

In both the total and LHCP modes (top and middle panels of Figure 4.10), there are several features that are not observed in the C/N_0 fade. For example, shortly after 1730 UT, there are two small impulses that are not evident in the

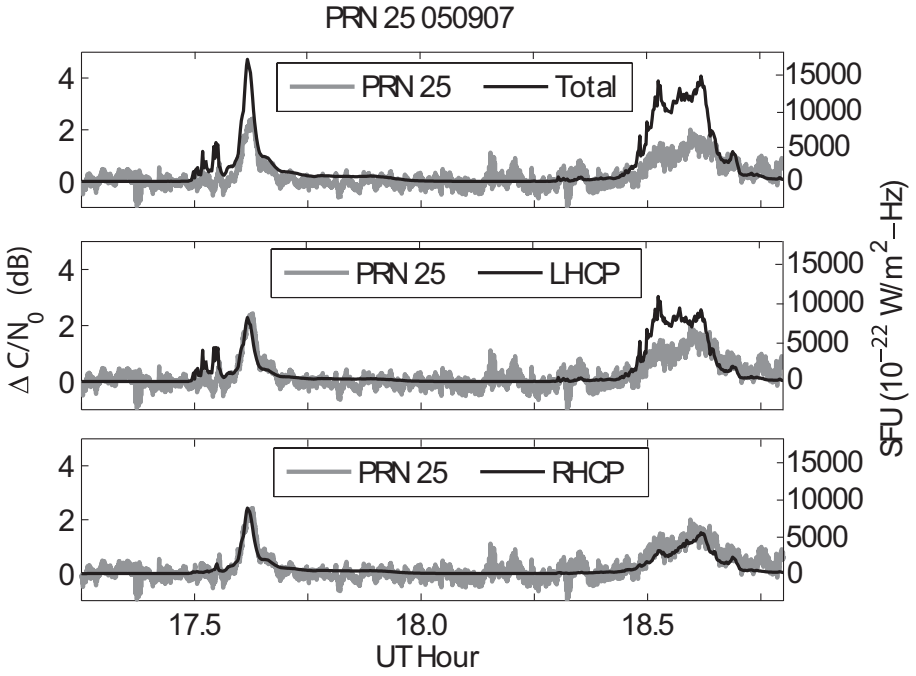

Figure 4.10: Total, LHCP, and RHCP mode solar radio burst power spectral densities (black) plotted against $\Delta C/N_0$ observed on PRN 25 (gray) as measured at Arecibo. The best association between the solar radio burst signal and $\Delta C/N_0$ is observed on the RHCP mode. The maximum C/N_0 fade at 1736 UT was 2.3 dB. The elevation of the sun at 1736 UT was 68° and at 1830 UT was 57°.

RHCP mode (bottom panel). If these two pulses in the LHCP mode had an effect on the GPS receiver, a corresponding increase in $\Delta C/N_0$ would be observed. Furthermore, if the receiver were sensitive to the LHCP mode, the C/N_0 fade seen at 1830 would have closely followed the shape of the LHCP mode solar radio burst. Clearly, there is a strong association between the C/N_0 fade and the RHCP mode. This result is expected since the GPS system utilizes only the RHCP mode for signal transmission and reception, and the antenna used for acquiring these data is only sensitive to the RHCP mode. Small discrepancies may still exist because the change in antenna gain as the sun changes elevation is not accounted for in these comparative plots.

Figure 4.11 depicts the measured C/N_0 fade for satellites at various elevations

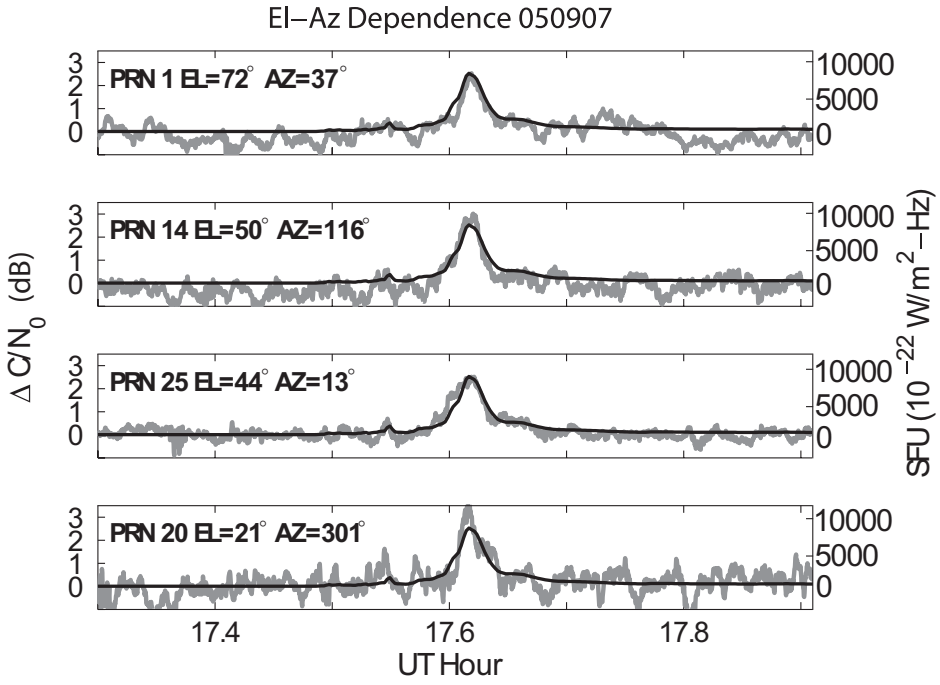

Figure 4.11: Carrier-to-noise ratio fade for satellites at various elevations and azimuths (gray) plotted against the RHCP solar radio burst power spectral density (black). The carrier-to-noise ratio fade (2.3 dB at L1) is nearly equal for all visible satellites. The magnitude of the fade was equal for all satellites, demonstrating the systematic effect of the solar radio burst on the receiver. The indicated elevation and azimuth of the satellites are at the peak intensity of the solar radio burst.

and azimuths (gray). Also plotted is the RHCP mode power spectral density of the solar radio burst (black). It is immediately apparent that the magnitude of the fades for all four satellites at 1736 UT are approximately equal. Hence, the solar radio burst has a systematic effect on the receivers and will affect all visible satellites regardless of elevation or azimuth.

The lack of elevation dependence of the fade also eliminates ionosphere *D*-region absorption as a cause of the C/N_0 fade. If the fades were a direct result of absorption at the L1 frequency (1.57542 GHz), then lower elevation satellites would exhibit deeper fades than the higher elevation satellites because of the different slant path. Furthermore, a 2.3 dB fade at L1 would require a *D*-region electron

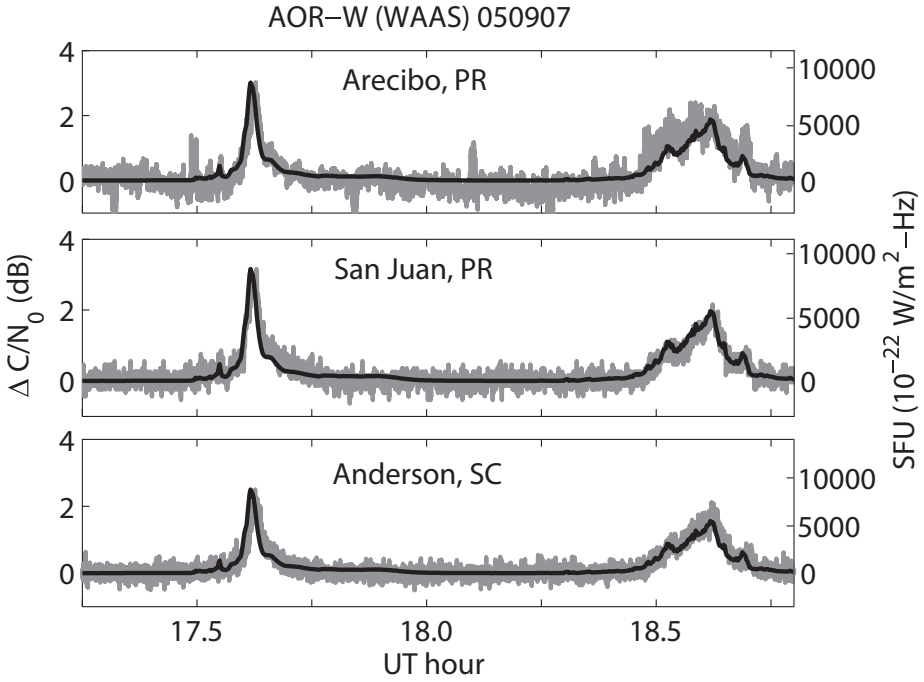

Figure 4.12: Carrier-to-noise ratio fade on the INMARSAT AOR-W (54°, PRN 122) WAAS satellite signal (gray) plotted against the measured solar radio burst for the 7 September 2005. The top panel is the C/N_0 fade as measured at Arecibo, PR using the Cornell ScintMon receiver. The middle and lower panels are the C/N_0 fades as measured using FAA WAAS receivers at San Juan, PR and Anderson, SC, respectively. Despite the geographic distances between the receivers, it is clear that all of the receivers are affected similarly by the solar radio burst. The elevation of the AOR-W WAAS satellite at Arecibo is 64.0°, 64.4° at San Juan, and 39.7° at Anderson. At 1736 UT, the elevation of the sun at San Juan and Anderson was 69° and 61°, respectively.

density of 10^{16} electrons/m^3, which is not physically possible.

To further eliminate potential local ionospheric irregularities or other terrestrial effects as a cause of the C/N_0 degradation, Wide-Area Augmentation System (WAAS) data from the Cornell ScintMon receivers were compared to WAAS data from Federal Aviation Administration (FAA) WAAS receivers located in San Juan, Puerto Rico (18.4°N, 66.0°W) and in Anderson, South Carolina, USA (34.5°N, 82.6°W).

WAAS was developed as an augmentation to the GPS system to provide ionospheric, ephemeris, and time corrections to single-frequency GPS users in near real-time. The corrections are made using measurements from 26 WAAS Reference Stations distributed throughout the United States of America. Each Reference Station is a NovAtel G-II Reference Receiver, with each unit containing two NovAtel L1/L2 GPSCard receivers and a MEDLL[2] (Multipath Estimating Delay Lock Loop) receiver (*Morrissey et al.*, 2002). Each receiver is phase locked to the same cesium frequency and utilizes a common antenna but operates autonomously and provides independent GPS measurements. Each receiver employs NovAtel's proprietary P-Code semi-codeless correlation technology, for which few details are available.

The measurements from each Reference Station are assimilated into a Kalman-filter estimator and the various corrections are transmitted to geostationary INMARSAT satellites, which retransmit the corrections on the GPS L1 frequency. The system is designed such that GPS receivers could utilize the WAAS signals without hardware modification, so the PRN sequences have the same period and number of chips as the GPS C/A codes. The navigation message, however, is significantly different. Regardless, the WAAS signal will suffer the same degradation as GPS signals due to solar radio bursts.

In Figure 4.12, the C/N_0 fade as recorded from the INMARSAT AOR-W (54°W, PRN 122) WAAS satellite as a result of the solar radio burst is shown for the receivers at Arecibo, PR; San Juan, PR; and Anderson SC in the top, middle, and lower panels, respectively. The carrier-to-noise ratio data for the FAA receivers are recorded at 1 Hz. In San Juan, the measured peak fade was 3.1

[2]The MEDLL is actually a series of NovAtel GPSCard receivers chained together to track each direct and reflected signal arriving at the antenna. See *Townsend et al.* (1995) for more details.

dB while at Anderson the peak fade was 2.5 dB. The discrepancies likely arise from differences in antenna gain that are not properly accounted for in the comparative plots. The physical distance between these receivers eliminate terrestrial (such as local radio frequency interference sources) or ionospheric effects (such as ionospheric irregularities causing scintillations) as the cause of the C/N_0 fades.

4.3.2 L2 Receiver Response

The impact of the solar radio burst on semi-codeless receivers was analyzed using data from an IGS receiver located in Christiansted, U.S. Virgin Islands (17.8°N 64.6°W) as well as from the FAA WAAS Reference Station located in San Juan, Puerto Rico.

The data from the IGS receiver in the U.S. Virgin Islands were collected using an ASHTECH Z-XII3 with a sampling rate of 1 Hz. As shown in Figure 4.13, $\Delta C/N_0$ at L1 (top panel, gray) is in good agreement with the solar radio burst power spectral density at 1.6 GHz (black) as reported by the OVSA. The fade at GPS L2 (middle and bottom panels, gray) is approximately twice that of the fade at GPS L1. In the bottom panel, the fade is not in good agreement with the solar radio burst power spectral density at either 1.2 GHz (black) or 1.6 GHz (red).

$\Delta C/N_0$, however, is in good agreement with the sum of the solar radio burst power spectral density at 1.2 GHz and 1.6 GHz, as shown by the black line in the middle panel. As discussed in Section 4.2.2, the fade at L2 is significantly deeper than the fade at L1 because the cross-correlating semi-codeless technique combines the additional noise power from the solar radio burst at L1 with the solar radio burst noise power at L2. Small discrepancies in the L2 carrier-to-noise ratio fade

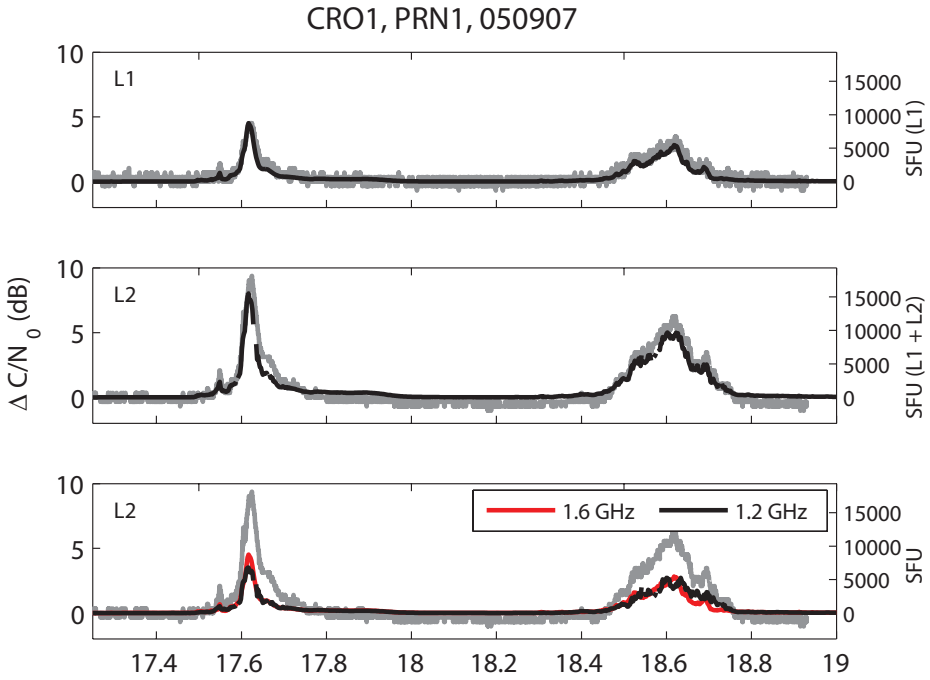

Figure 4.13: Carrier-to-noise ratio fade on GPS PRN 1 at L1 (top panel, gray line) and at L2 (middle and bottom panels, gray line). The data are acquired from the Christiansted, U.S. Virgin Island IGS receiver (17.8°N 64.6°W). The OVSA RHCP solar radio burst power spectral density at 1.6 GHz is represented by the black line in the top panels. The black line in the middle panel represents the sum of the solar radio flux at 1.2 and 1.6 GHz. In the bottom panel, the red line represents the 1.6 GHz OVSA RHCP solar radio flux, and the black line is the 1.2 GHz OVSA RHCP solar radio flux. All scales are equal. Note that the fade at L2 is roughly double the fade at L1. The noise power of the solar radio burst at 1.2 GHz and 1.6 GHz are nearly equal, and the semi-codeless tracking technique effectively sums the noise power at the two frequencies, as shown in the middle panel.

and the sum of the solar radio burst power spectral density at 1.2 and 1.6 GHz can be attributed to differences in antenna gains at the L1 and L2 frequencies. The gain parameters for the antennas are not know, so the scaling may be biased.

The data from the San Juan FAA WAAS G-II receiver are also analyzed (Figure 4.14), and surprisingly the fade at L2 was not deeper than the fade at L1, as is shown for the Christiansted receiver (Figure 4.13). Differences in antenna gain

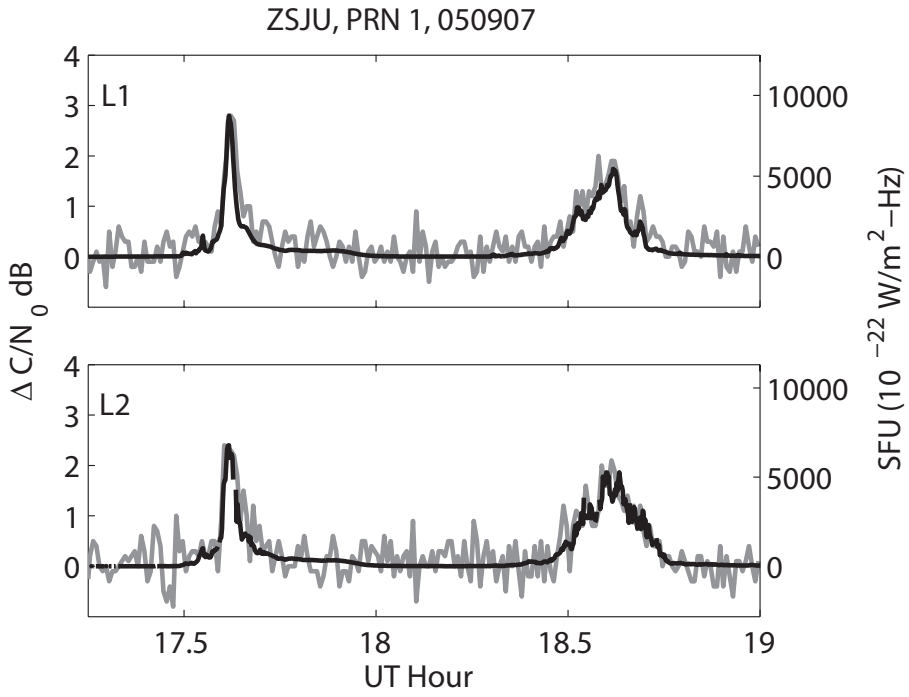

Figure 4.14: Carrier-to-noise ratio fade on GPS PRN 1 at L1 (top, gray) and at L2 (bottom, gray). The data are acquired from the San Juan, Puerto Rico FAA WAAS reference station at a cadence of 30 seconds. The RHCP solar radio burst power spectral density at 1.6 GHz is represented by the black line in the top plot, and also at 1.2 GHz by the black line in the bottom plot. The axis on the left is equal for both plots, but the plot on the right, corresponding to the solar radio burst intensity, differs.

from L1 to L2 cannot account for this discrepancy because the fades at L2 should be at least as large as the fades at L1. Apparently the WAAS G-II receiver reports the true carrier-to-noise ratio at L2, as though the receiver were utilizing the encrypted military codes. Inquiries were made to both NovAtel and individuals at Zeta Associates Inc.[3] on the operation of the receiver. Unfortunately the semi-codeless tracking techniques and recovery of carrier-to-noise ratio at L2 are NovAtel proprietary. It is beyond the scope of this thesis to conjecture how the WAAS G-II receiver is capable of producing these results.

[3]Zeta Associates Inc. is a key provider of WAAS Reference Station and signal integrity monitoring.

The 7 September 2005 solar radio burst is not a particularly devastating event. Only in rare cases would a receiver be significantly impacted by the solar radio burst (e.g. a receiver that is operating in extremely weak signal environments may suffer from loss of lock or loss of navigation solution). GPS receivers can typically navigate with carrier-to-noise ratios greater than 25 dB-Hz. A 3 dB fade is not large enough to degrade tracking loop performance, which in turn would degrade ranging measurements and position estimates. The event is important, however, because these are the first direct observations of a solar radio bursts' impact on a GPS receiver, and they allow for verification of the computational analysis in Section 4.2.

4.4 Further Space Weather Impacts Associated with Solar Radio Bursts

Solar radio bursts are often accompanied by associated solar flares and sometimes ensuing geomagnetic storms. X-ray flares provide more energy for the ionization of neutral gasses in all regions of the ionosphere. Intense solar flares, such as the 7 September 2005 event, can penetrate well into the *E*- and *D*-regions, causing significant absorption of high-frequency radio (<50 MHz) signals. The increase in electron density also causes a phase advance of the GPS carrier signals, and an equal but opposite delay of the PRN codes as they propagate through the ionosphere. Using dual frequency receivers, the difference in arrival time of the PRN codes on L1 and L2 can be used to infer the ionospheric electron density.

An Ashtech MicroZ-CGRS dual-frequency receiver is located at the Arecibo Observatory for obtaining Total Electron Content (TEC) measurements. TEC is

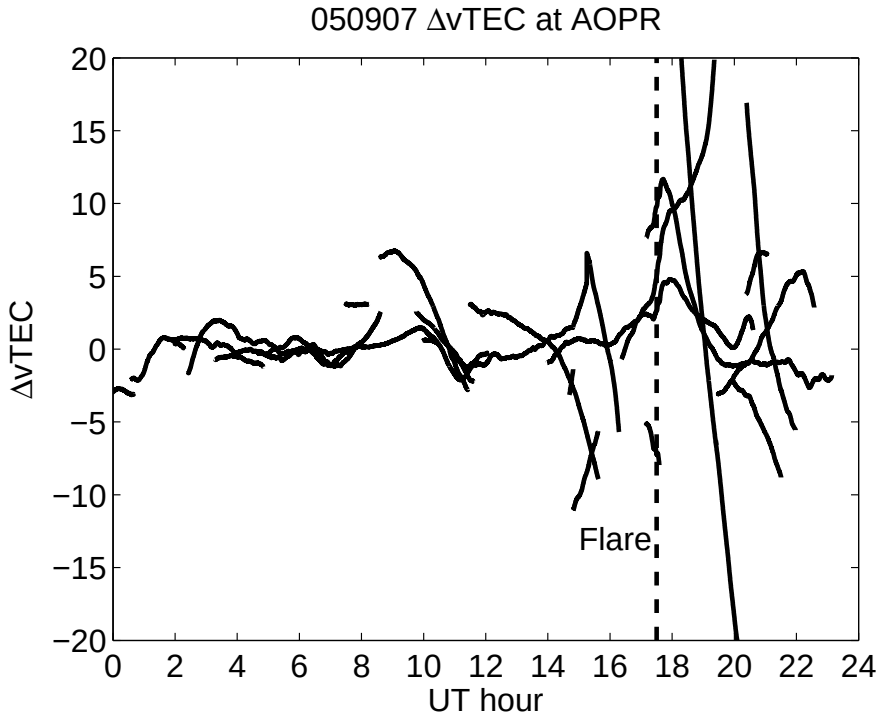

Figure 4.15: Each black trace represents the difference in Vertical Total Electron Content (ΔvTEC) for a satellite from 7 September 2005 to 6 September 2005. The data are recorded at Arecibo Observatory, Puerto Rico. Only observations above 15° elevation are used. vTEC increased by 3-5 TEC units for the duration of the flare.

the measure of the total number of electrons in a 1-m^2 column between the receiver and the satellite. The Ashtech receiver is ideally suited for intense ionospheric activity as it has special-purpose hardware and tracking loops to avoid the effects of a rapidly changing signal environment. The receiver makes 10 sample per second line-of-sight TEC measurements that are then processed for vertical TEC (vTEC). vTEC is a measure of the total number of electrons in a 1-m^2 column at the 350 km satellite pierce point derived from the line-of-sight TEC via an obliquity factor.

In Figure 4.15, the difference in vertical TEC (ΔvTEC) from 7 September 2005 as compared to 6 September 2005 at Arecibo Observatory is plotted as a function of local time for all visible satellites above 15° elevation during the day. At 17.6

UT hours, there is an increase in the vertical TEC of about 3-5 TEC units (1 TEC unit = 10^{16}electrons/m^2), corresponding to the increased ionization in the ionosphere due to the solar flare. The results are in agreement with other events, including the Halloween events of 2003 (*Tsurutani et al.*, 2005).

An increase of 3-5 TEC units corresponds to an increased ranging error of about 50 - 80 cm at zenith. For low elevation satellites, the delay typically increases by a factor of three. For a single-frequency (L1-only) user, the increased ranging error maps into the navigation solution error as a function of the Dilution of Precision. For the L1-only user, the increase in ranging error may actually substantially increase or decrease the overall system error because it is likely that the transmitted ionospheric corrections misrepresent the true state of the ionosphere (see Section 2.3.3). A dual-frequency receiver, on the other hand, is capable of eliminating the error induced by the sudden increase in electron density. It should be noted that increases of 3 - 5 TEC units correspond to phase changes of about 2.5 - 4 cycles at L1, and about 2 - 3.3 cycles at L2. The change occurs over several minutes. The fastest rate of change in TEC is just about 1.17 TEC units (1 cycle at L1) over roughly three minutes. These rates of change in carrier phase are not rapid enough to create loss loss of carrier lock on receivers, contrary to the results suggested by *Chen et al.* (2005).

In addition to the TEC measurements, beginning at about 17:48 UT (about 8 minutes after the onset of the flare), the Arecibo Observatory 430 MHz line-feed radar began collecting vertical electron height profiles. The data are available from a minimum height of about 59.1 km to a maximum height of 531.9 km with a 300 meter range resolution. The data are collected using a 13-baud Barker code $(+ + + + + - - + + - + - +)$ with a baud length (T_B) of 4 μs. Hence, the

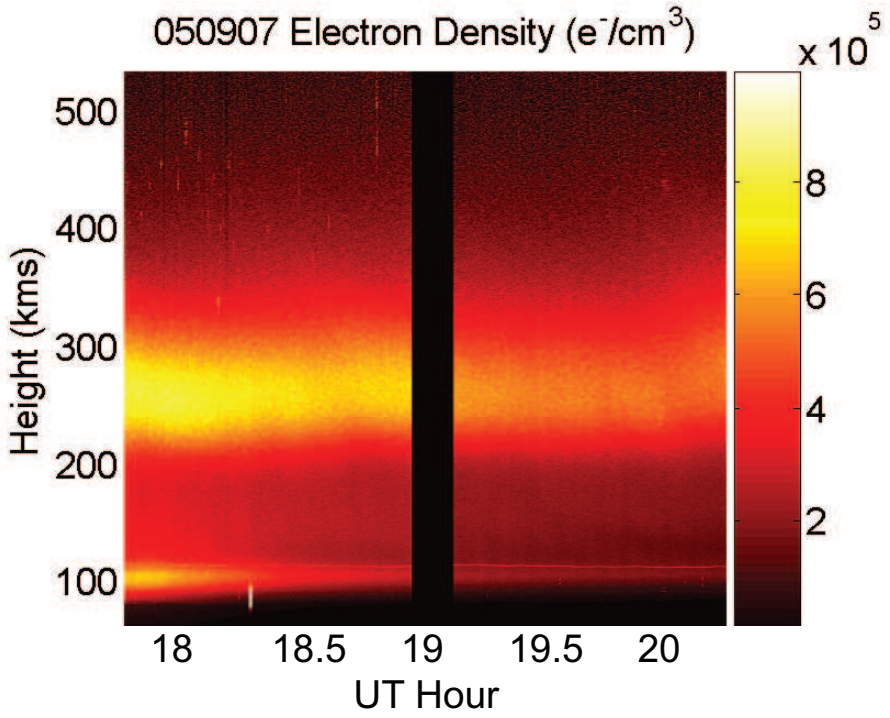

Figure 4.16: Electron density as a function of height using the Arecibo Incoherent Scatter Radar. The radar began acquiring data at 17:48 UT (approximately 8 minutes after the onset of the flare). Apparent in the figure is the intense ionization in the D- and E-regions of the ionosphere, which immediately recombines as the x-ray and EUV flux from the flare decreases.

total pulse length was 52 μs. The inter-pulse period (IPP) was 10 ms, and the data are integrated over 20 seconds. No near-field corrections were applied, so the electron densities represented in Figure 4.16 are under-estimated for altitudes less than about 200 kilometers.

A distinct increase in the overall electron density extending from the *D*-region (70-90 kilometers), through the *E*-region (90 - 150 kilometers), and through the *F*-region is evident. The largest increase in density, as expected, occurred in the *D*-region and *E*-region, where the vast majority of solar ultra-violet (UV) and x-ray flux are absorbed. Once the flare dissipates around 19:00 UT, the intense *D*- and *E*-region ionization has subsided entirely. Furthermore, there is evidence of a

sporadic-E layer, but no other irregularities. With the exception of the increase in electron density, the ionosphere above Arecibo is quiet, so irregularities are not a cause of the signal fade.

4.5 Conclusion

In this chapter, the effects of solar radio bursts on the Global Positioning System have been detailed. A more realistic computational analysis derived from *Klobuchar et al.* (1999) is shown to determine the size of a solar radio burst that would cause a 3 dB drop in received carrier-to-noise ratio at L1. The expected fade was also derived for the L2 signal for both coded and semi-codeless receivers. The receiver's sensitivity to carrier-to-noise ratio was also analyzed to determine when they are expected to fail. The results were compared to the relatively modest solar radio burst that occurred on 7 September 2005 and shown to be consistent with the computational analysis. The 7 September 2005 events also show that the solar radio burst affects all receivers in the sunlit hemisphere, and that all satellites tracked by the receiver are affected equally. Finally, *D*-region absorption and ionospheric scintillation is eliminated as a potential cause of the signal fade recorded on the sunlit receivers.

The observations and computational analysis allow for predictions. *Nita et al.* (2002) conducted a statistical study of the peak flux distribution of 40 years of solar radio burst data as a function of frequency and time. Their results can be used to estimate how often a solar radio burst of a certain magnitude at a given frequency will occur. When the correction factor for missed observations is applied, a solar radio burst of at least 20,000 SFU (total polarization) in the 1.0 - 1.7 GHz

range will occur every 244 days over a solar cycle. During solar maximum, the time interval decreases to about 86 days. For the noise figures and antenna gains used in the computational analysis, solar radio bursts of 20,000 SFU imply fades of 3 - 6 dB depending on the polarization of the burst.

However, solar radio bursts on record (measured at 1.415 GHz) have exceeded 40,000 SFU total peak power at least 14 times since 1967 (*Klobuchar et al.*, 1999). From the National Oceanographic and Atmospheric Administration database, the most energetic occurred in April 1973 and had a total peak intensity of 165,000 SFU at 1415 MHz. The second largest occurred on 16 February 1979 and had a total peak intensity of 88,000 SFU (*Klobuchar et al.*, 1999). This particular burst sustained at least 40,000 SFU for 72 minutes.

Since the magnetic fields surrounding the site of the solar radio burst have equal probability of polarization, at least 50% of the power is expected to arrive in the RHCP mode for at least 50% of the duration of the solar radio burst. A solar radio burst of 40,000 - 80,000 SFU RHCP would imply a 9 - 12 dB C/N_0 fade at L1. For inadequately designed systems, the fades on L1 could cause loss of lock on all satellites. Failure of the receiver to produce a navigation solution is entirely possible during 80,000 SFU (total polarization) solar radio bursts. Semi-codeless L1/L2 receivers are expected to fail. As will be shown in Chapter 5, the future came sooner than expected.

Acknowledgements

TEC data for 6 and 7 September 2005 provided by Dr. Bill Rideout and Dr. Anthea Coster of the MIT Haystack Observatory, in collaboration with the National Astronomy and Ionosphere Center (NAIC) at the Arecibo Observatory. Arecibo

radar data courtesy of Nestor Aponte and Hien Vo. Dale Gary of the New Jersey Institute of Technology provided the OVSA data for the September 2005. 1-second WAAS data were provided by Tom McHugh of the FAA.

Chapter 5

Record Setting December 2006 Events

Prediction is very difficult, especially if it's about the future. - Neils Bohr

5.1 Introduction

Solar radio bursts during December 2006 were sufficiently intense to be measurable with GPS receivers. The strongest event occurred on 6 December 2006 and was coincident with an X6 solar flare from active region 10930 (Figure 5.1). The solar flare triggered a solar radio burst that affected the operation of many GPS receivers. This event exceeded 1,000,000 SFU and was about ten times larger than any previously reported event. The strength of the event was especially surprising since the solar radio bursts occurred near solar minimum. The strongest periods of solar radio burst activity lasted several minutes. The events caused wide-spread outages of GPS receivers in the sunlit hemisphere of the Earth. Receivers that did not suffer from outages suffered from increased positioning errors during the most intense periods of the solar radio burst. Other events on 13 December 2006 and 14 December 2006, although not quite as energetic as the 6 December event, also shattered previous records.

The unprecedented size and timing (they occurred at solar minimum) of these solar radio bursts came as a complete surprise to many. Indeed, the effects of the solar radio burst were widespread. In the subsequent sections, the effects of the

*Portions of this chapter are based on the publication *Cerruti et al.* (2007), which is joint collaborative work with A. P. Cerruti, P. M. Kintner, D. E. Gary, A. J. Mannucci, R. F. Meyer, P. Doherty, and A. J. Coster. Reproduced by permission of the American Geophysical Union.

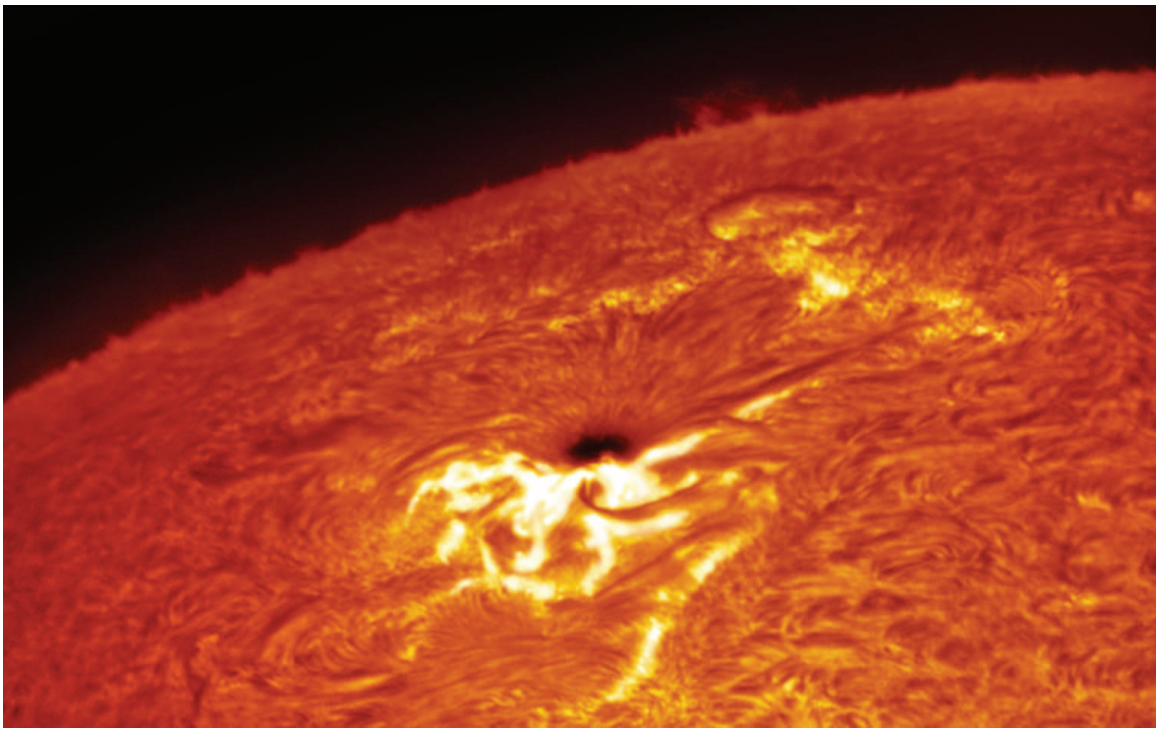

Figure 5.1: Hydrogen-alpha photograph of sunspot 930 during the 6 December 2006 X6 solar flare. This solar flare generated the record-setting solar radio burst that significantly impacted GPS receivers. Figure © 2006, Gary Palmer (http://theSuninmotion.com/). Reproduced with permission.

solar radio bursts on various Global Positioning Systems are analyzed.

5.2 6 December 2006 Solar Radio Burst

The December 2006 solar radio burst events were unexpected because they occurred near solar minimum, and yet these events proved to be record-setting. The sunspot region responsible for these events was active region 10930. Initially, on 5 December 2006, the sunspot group unleashed an X8 solar flare that was not accompanied by a solar radio burst at the GPS frequencies. On 6 December, however, the region erupted again with an X6 solar flare that produced an extremely energetic solar radio burst in the GPS frequencies. According to *Cerruti et al.*

Table 5.1: Peak right-hand circularly polarized flux densities as measured by the Owens Valley Solar Array (OVSA) and the FST. Both instruments are located in Owens Valley, California and operated by the New Jersey Institute of Technology (NJIT). The OVSA has a large frequency range (1.2 - 18 GHz), but is limited in its sampling time. The FST, however, samples continuously but is limited to frequencies from 1.0 - 1.5 GHz. Both instruments exhibited saturation effects. Data courtesy of Dale Gary, NJIT.

Instrument	Frequency (GHz)	Peak Flux Density (SFU)	Range (SFU)
OVSA	1.2	$6.5 \cdot 10^5$	$4.0 \cdot 10^5 - 8.5 \cdot 10^5$
	1.4	$1.0 \cdot 10^6$	$0.48 \cdot 10^6 - 1.1 \cdot 10^6$
	1.6	$0.5 \cdot 10^6$	$0.43 \cdot 10^6 - 0.93 \cdot 10^6$
FST	1.227	$1.85 \cdot 10^6$	n/a
	1.405	$2.51 \cdot 10^6$	n/a

(2007), the solar radio burst produced 10 times more radio noise than the previous record and produced about 20,000 times more radio noise than the entire rest of the Sun. The solar radio burst was particularly energetic because conditions above the sunspot group were conducive for natural maser emissions (*Treumann*, 2006). The solar radio burst consisted of many impulsive solar radio bursts of 3 - 4 MHz bandwidth that, in some cases overlapped (*Cerruti et al.*, 2007). The solar radio burst lasted over an hour, with the most intense emissions from about 1930 UT to about 1938 UT.

In Figure 5.2, the power spectral density of the RHCP solar radio burst from 1.2 - 3 GHz is plotted from 1839 UT to 2000 UT. In Figure 5.3, the LHCP solar radio burst power spectral density is also plotted using the same time, frequency, and solar flux unit scales. The solar radio burst is strongly RHCP for the duration of the event, especially at the GPS L1 frequencies. The highly energetic impulsive events ($\sim 10^5$ SFU) on the LHCP plot at 1.2 GHz are likely due to cross-talk between the RHCP and LHCP modes.

Exact determination of the peak flux density is problematic primarily because of saturation effects on solar radio burst measuring equipment. The receivers are equipped with programmable attenuators that automatically switch in up to 55 dB of attenuation, but the flux density of the burst varied so rapidly at times that the automatic switching could not keep up with the variations. Measurements on different antennas show differences that are due to non-linearity caused by compression of the second-stage amplifier. Therefore, the absolute calibration of the burst is uncertain, but it is believed that the flux density exceeded 10^6 SFU. The best estimates from the OVSA are summarized in Table 5.1, along with flux density estimates from the Frequency Agile Solar Radiotelescope (FASR) Subsystem Testbed (FST).

The FST is a three-element interferometer collocated with the OVSA was operating at the same time to provide 1 MHz frequency resolution and 20 ms time resolution during the burst (*Liu et al.*, 2007). This system also suffers from compression of the second-stage amplifier during the strongest period of the burst, but can be cross-calibrated with OVSA to provide detailed information about the emission mechanism. The FST provides a nearly continuous measurement (20 ms sampling rate) of the solar radio burst from 1 - 1.5 GHz, while the OVSA takes measurements from two antennas, at discrete frequencies, every 8.1 seconds. In addition to the saturation effects, the burst is quite dynamic temporally, and the data collection method of each instrument is different, so the peak values differ substantially.

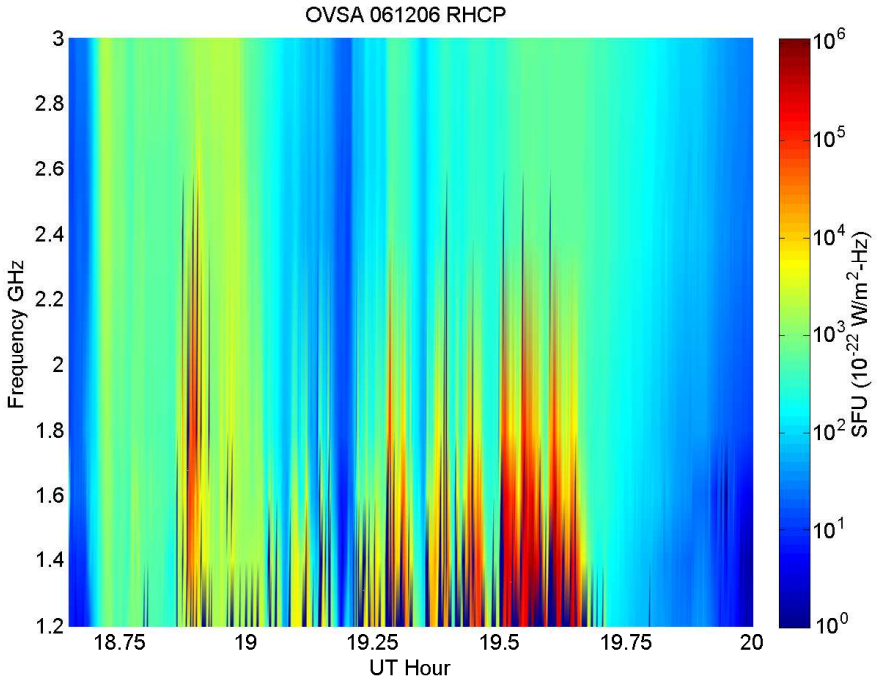

Figure 5.2: Owens Valley Solar Array RHCP measurements of the 6 December 2006 solar radio burst that accompanied the X6 solar flare from active region 10930. Rapid changes of solar radio burst power density required gain changes in the instrument, so some measurements are missing and have been interpolated. The solar radio burst was highly energetic in the 1 - 2 GHz range, but above 3 GHz the power density rarely exceeded 10,000 SFU. The burst is clearly composed of many impulsive events.

5.2.1 FAA WAAS System Response

The intensity of the solar radio burst on 6 December 2006 is notable on the WAAS system, but it did not disable the system. In Figure 5.4 the carrier-to-noise ratio as recorded for PRN 17 on the L1 and L2 frequencies at the FAA WAAS reference receiver located in Houston, Texas (29.6°N, 95.2°W) is shown. The peak fade at L1 is approximately 17.5 dB, corresponding to a minimum carrier-to-noise ratio of 34.5 dB-Hz. The peak fade at L2 is approximately 20.0 dB, corresponding to a minimum carrier-to-noise ratio of 24.7 dB-Hz. The standard deviation of C/N_0 at the peak of the fade is approximately 0.69 dB at L1 and larger than 0.58 dB at

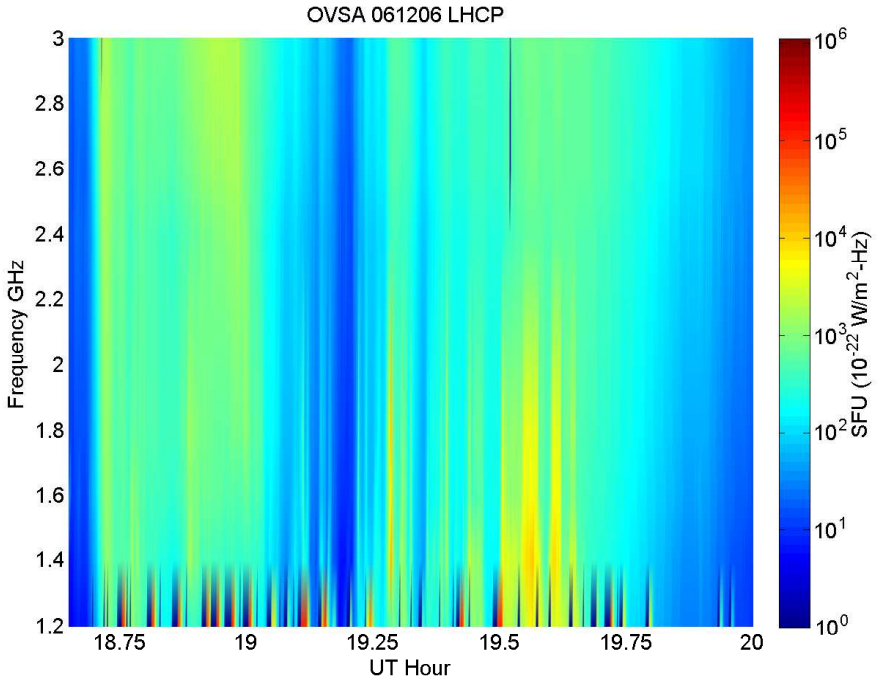

Figure 5.3: Owens Valley Solar Array LHCP measurements of the 6 December 2006 solar radio burst that accompanied the X6 solar flare from active region 10930. Rapid changes of solar radio burst power density required gain changes in the instrument, so some measurements are missing and have been interpolated. Most of the solar radio burst power spectral density is in the RHCP mode (see Figure 5.2) by at least a factor of ten. The time, frequency, and solar flux unit scale are the same as in Figure 5.2.

L2[1]. At the time of the peak fade, the sun was at an elevation of approximately 34.2°.

An estimate of the peak fade observed on the WAAS receivers to the maximum solar radio burst power spectral density can be made based on Equation 6.20. At an elevation of approximately 34.2° at the L1 frequency, the gain of the antenna used on the WAAS receiver is about -0.18 dBic. The noise figure of the WAAS Reference Stations is determined by the low-noise amplifier in the antenna, and

[1]The standard deviation of C/N_0 at L2 was calculated for a nominal carrier-to-noise ratio of 34.1 dB-Hz. Data at 24.7 dB-Hz was insufficient to calculate the standard deviation. The actual value at L2 is probably significantly higher than the reported value, since the standard deviation of C/N_0 increases as the carrier-to-noise ratio falls.

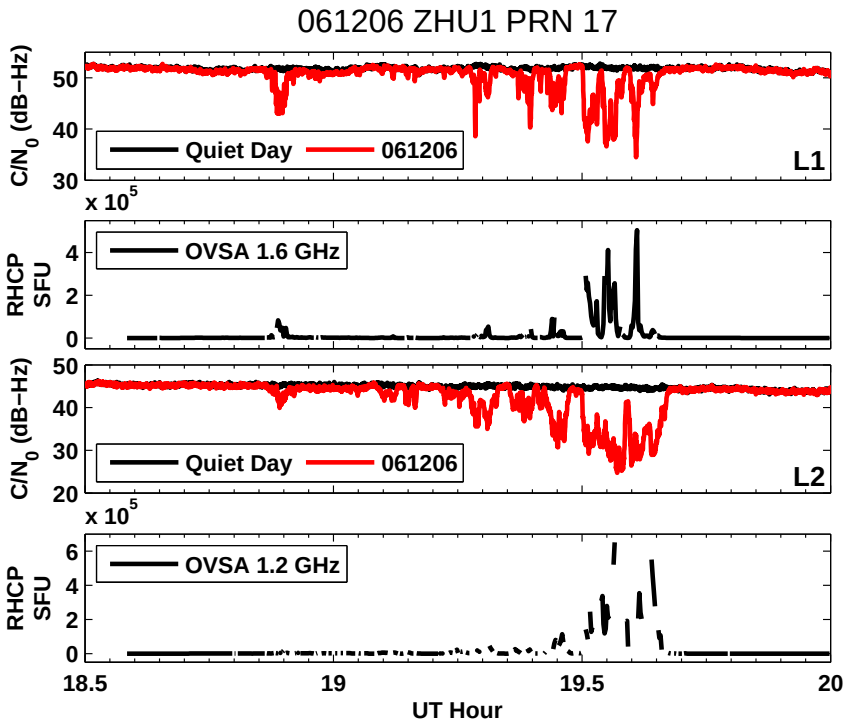

Figure 5.4: Comparison of the carrier-to-noise ratio fade at L1 (first panel) on 6 December 2006 (red) and the day prior to the solar radio burst (black). The solar radio flux at 1.6 GHz as measured at the OVSA is shown in the second panel. The fade at L2 is shown in the third panel, and the OVSA solar radio flux at 1.2 GHz is shown in the fourth panel.

is better than 2 dB (private communication with Patrick Reddan). The ambient temperature at Houston was 17 °C, and assuming a sky noise of approximately 100 K, the 17.5 ± 0.69 dB fade at L1 at 1936 UT will occur for a RHCP solar radio burst of about 633,000 - 874,000 solar flux units RHCP. At L2, the gain of the antenna is -3.44 dBic, so for the 20 ± 0.58 dB fade at 1934 UT, a RHCP solar radio burst of about 1,499,000 - 1,963,000 solar flux units is required. At 1.6 GHz, the OVSA reported a peak RHCP solar radio burst of 500,000 SFU, and at 1.2 GHz, a peak value of 650,000 SFU. The FST, on the other hand, reported 1,850,000 SFU RHCP at 1.227 GHz and about 2,510,000 SFU at 1.405 GHz.

Figure 5.5 illustrates the carrier-to-noise ratio fades as observed at the WAAS

Reference Stations at various points during the solar radio burst event. The color of the dot represents the intensity of the fade, as indicated by the color scale. The location of each dot represents the position of the WAAS Reference Station. Only 15 WAAS Reference Stations are indicated because carrier-to-noise ratio data are not available for all 26 Stations. Note that the WAAS Reference Stations located in darkness are not affected by the solar radio burst, where as the sunlit receivers are all affected. At the peak of the solar radio burst (Panel C), the response of the receivers is fairly consistent, with a tendency of receivers closer to the subsolar point to exhibit a larger fade as a result of the increased antenna gain at higher solar elevation angles.

The numbers are in good agreement with the OVSA and FST reported values, but discrepancies can arise for several reasons. First, the carrier-to-noise ratio from the WAAS receivers is reported every second. It is not known if the 1-second carrier-to-noise ratio is an average over the second, or if it represents an arbitrary carrier-to-noise ratio for an integration period (typically 20 ms) over the previous second. Furthermore, the bandwidth of the GPS receivers is approximately 18 MHz at both L1 and L2. The exact noise figures of the receiver are difficult to ascertain accurately and can vary depending on the ambient temperature of the receiver and amplifiers. The measurement bandwidth is substantially narrower for the FST (1 MHz) and significantly larger for the OVSA (2 side-bands of 50 MHz each) and both are at different frequencies than GPS L1 and L2. Finally, the OVSA exhibited saturation effects while measuring the solar radio burst. Regardless, the fades seen on the GPS data are consistent with magnitude of the fade measured on the OVSA. A better estimate of the solar radio burst power as derived from GPS measurements is detailed in Chapter 6, Section 6.3.

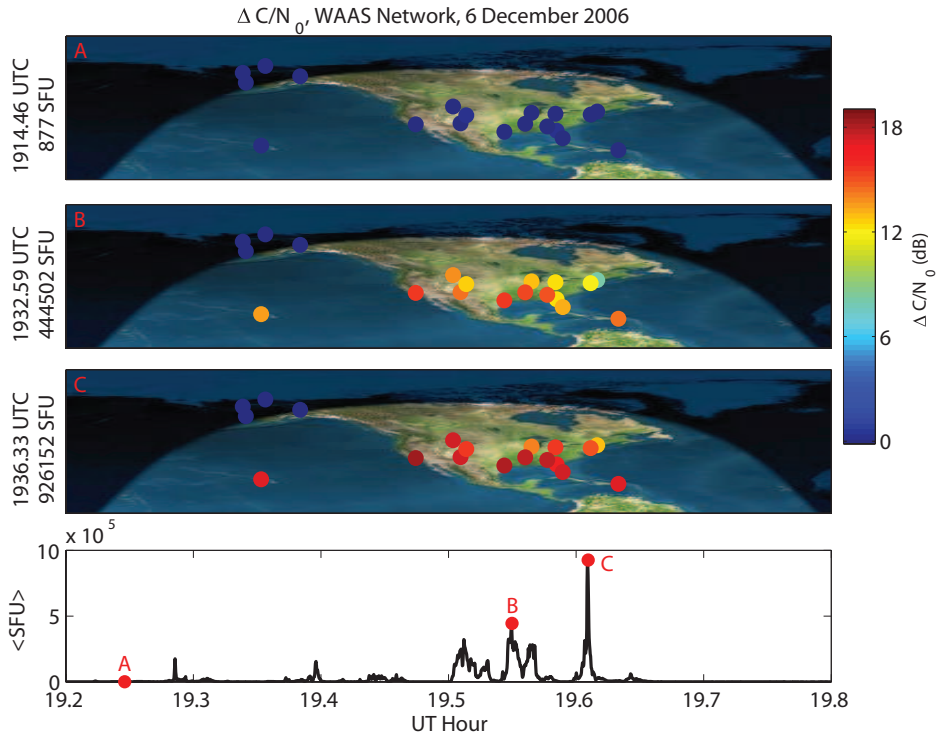

Figure 5.5: Measure of fading observed over the WAAS Reference Station network. Panel A illustrates that the measured fade prior to the solar radio burst is near zero. Panel B and Panel C illustrate the observed fades during the solar radio burst at particular times, noted to the left of each plot. Each colored dot represents the fade observed at that Reference Station. The bottom panel illustrates the solar radio burst RHCP power derived from the measured fades at each point in time (detailed in Chapter 6, Section 6.3). Note that in general, the Stations closest to the subsolar point exhibit a larger fade, as expected. The subsolar point lies due south of Mexico in the southern hemisphere. Image of Earth © The Living Earth, 1996 and is used here by permission of the publisher. Day/night overlay created using Earth Viewer by John Walker.

Despite the large fades observed, no WAAS receiver failed to navigate (each receiver was tracking four or more satellites) over the duration of the solar radio burst. According to *Cerruti et al.* (2007), the WAAS Miami reference receiver was most impacted by the solar radio burst. For a period of 1.5 min the receiver was only tracking four satellites. All other receivers in the network were tracking five or more satellites. Over the duration of the solar radio burst, the stability of the WAAS GPS equipment enabled the continuous operation of the en-route (Non-

Precision Approach) level of service throughout the event. The system, however, was not totally unaffected. The vertical guided approach service (LPV), used for landing aircraft, was impacted for a period of less than 15 minutes over the northwest United States. This is the first instance of a solar radio burst having an impact on the WAAS system. The WAAS system, however, suffers from periods of loss of service of varying intensity and duration from events other than solar radio bursts including geomagnetic storms and satellite maintenance. Redundancy and safeguards are built into the system to cope with space weather events and other expected and unexpected scenarios (e.g. satellite failures).

Each panel of Figures 5.6-5.8 depicts the total number of tracked satellites at each of 27 WAAS Reference Stations. The numbers quoted here differ slightly than those from *Cerruti et al.* (2007) because only data from the first of three receivers located in each Reference Station are utilized. The bottom panel in each figure shows the OVSA RHCP solar radio burst power spectral density. The panels are organized by increasing solar elevation angle.

A movie (WAAS_061206.mov) depicting the number of observations over time can be found on the accompanying CD. In the movie, each Reference Station is plotted as a dot, where the color of the dot indicates the number of dual-frequency observations made from each satellite according to the color scale on the right. The OVSA solar radio burst flux density at 1.6 GHz is shown in the bottom panel.

Since the WAAS system depends on dual-frequency measurements to determine the ionospheric corrections, the figures and movie show the number of dual-frequency code and carrier phase observations available to each receiver. In other words, for each satellite tracked, if the L1 and L2 code and carrier phase observations are valid, it is considered an observation. An analysis was not performed

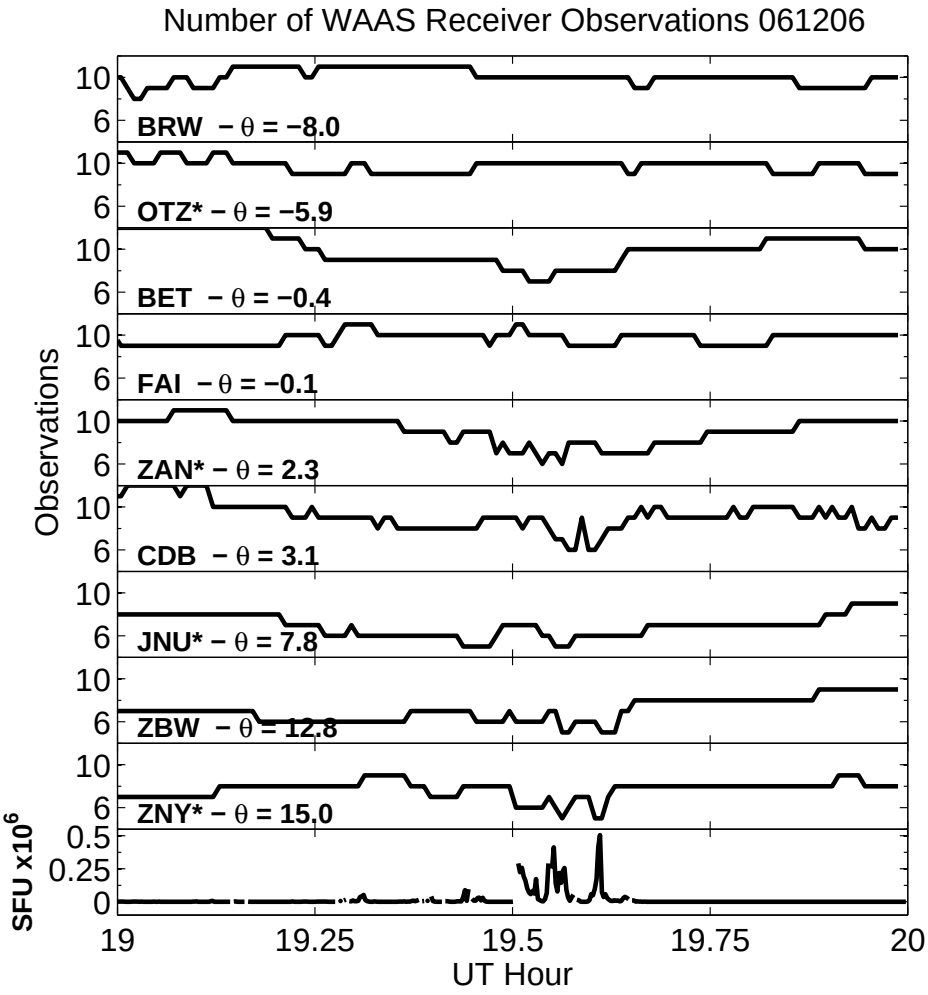

Figure 5.6: Number of satellites tracked for several WAAS Reference Receivers. The sites are organized by the solar elevation angle at 1936 UT. A satellite is not considered to be tracked if there is a complete loss of code or carrier phase measurements on either L1 or L2. Carrier phase cycle slips are not considered. Sites denoted with a * are in darkness at the time of the solar radio burst. Reference Station JNU (Juneau, AK) is marked with a * because, although the solar elevation angle was positive, the receiver did not suffer drastically from the effects of the solar radio burst. Perhaps the sun was behind a mountain or a large obstacle at the time of the solar radio bursts. θ is the solar elevation angle at each Station. The OVSA RHCP solar radio burst power spectral density at 1.6 GHz is plotted in the bottom panel for comparison.

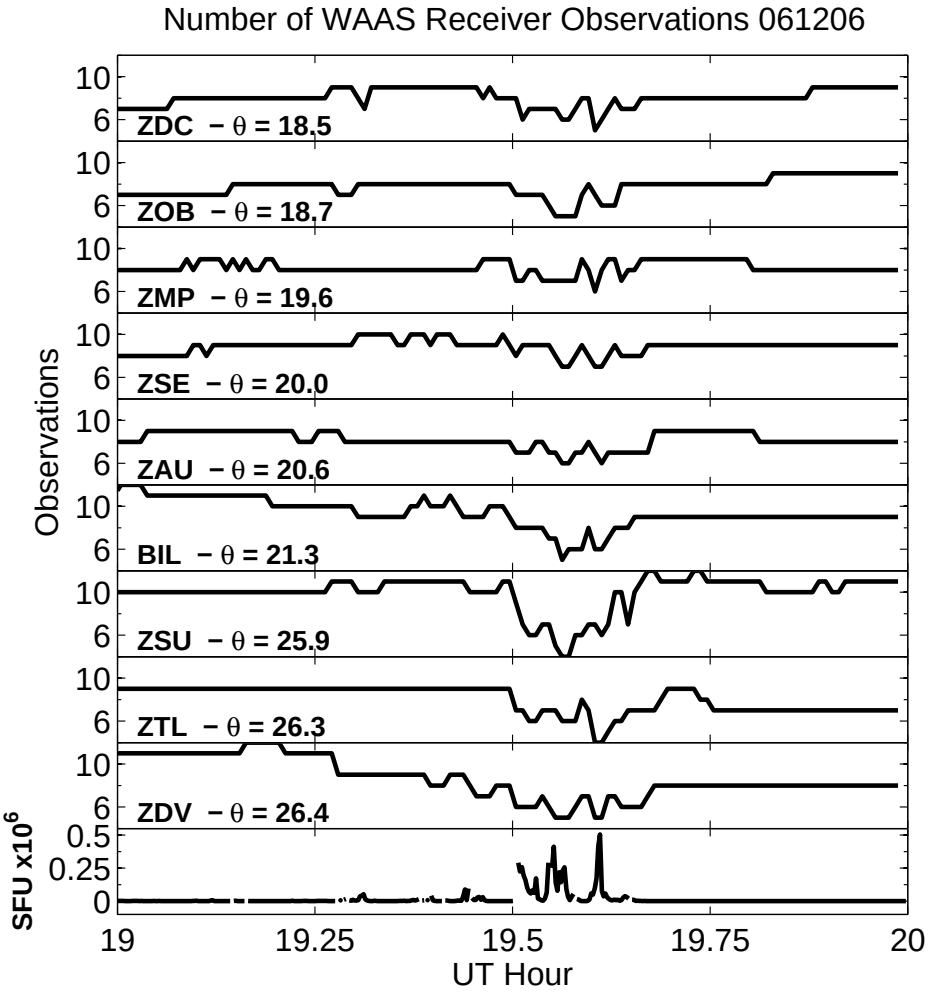

Figure 5.7: Number of satellites tracked for several WAAS Reference Receivers. The sites are organized by the solar elevation angle at 1937 UT. A satellite is not considered to be tracked if there is a complete loss of code or carrier phase measurements on either L1 or L2. Carrier phase cycle slips are not considered. θ is the solar elevation angle at each Station. The OVSA RHCP solar radio burst power spectral density at 1.6 GHz is plotted in the bottom panel for comparison.

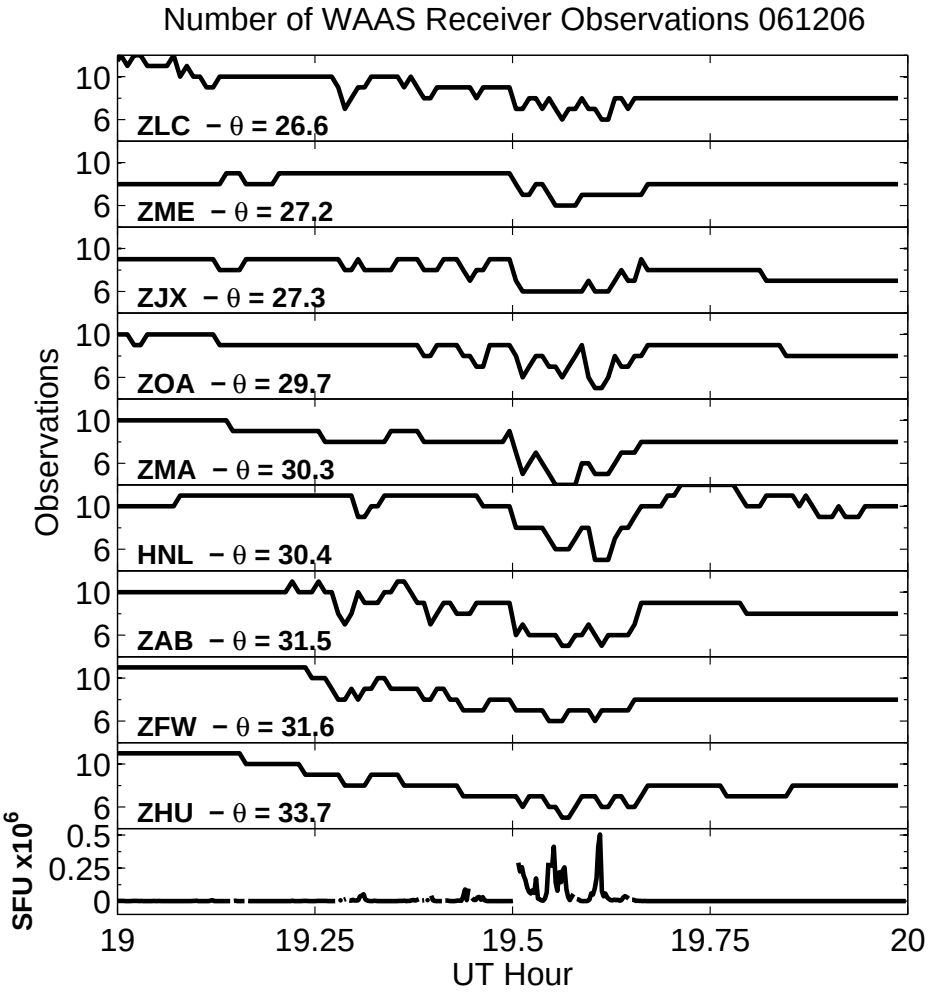

Figure 5.8: Number of satellites tracked for several WAAS Reference Receivers. The sites are organized by the solar elevation angle at 1937 UT. A satellite is not considered to be tracked if there is a complete loss of code or carrier phase measurements on either L1 or L2. Carrier phase cycle slips are not considered. θ is the solar elevation angle at each Station. The OVSA RHCP solar radio burst power spectral density at 1.6 GHz is plotted in the bottom panel for comparison.

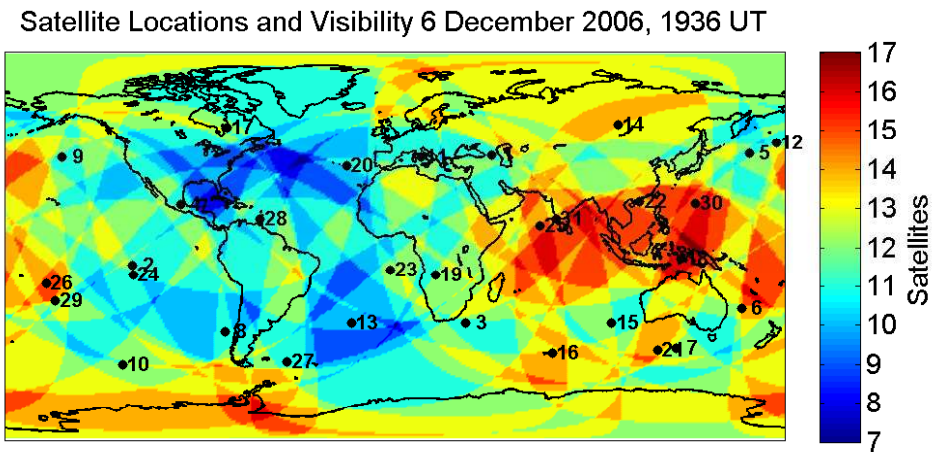

Figure 5.9: Number of satellites above 0° elevation on 6 December 2006 at 1936 UT. At this time, there are 8 - 11 visible satellites over the continental United States.

to determine if cycle slips occurred in the carrier phase observations. Cycle slips should be considered as they are a loss of carrier lock on the satellite that will render TEC measurements invalid if not detected and accounted for. The WAAS receivers provide a robust algorithm for detecting cycle slips.

In general, the number of observations at each WAAS Reference Station decreases with increasing solar elevation angle. The gain pattern of the WAAS receiving antennas is responsible for this phenomenon, since higher gain occurs at higher elevation angles. Some receivers show less effect than others, and this may be due to obscuration of the sun by nearby obstacles such as buildings or mountains (e.g. Juneau - JNU). Furthermore, some receivers may seem more susceptible than others (e.g. Atlanta - ZTL, or San Juan - SJU) since they are tracking only 4 satellites, but in reality the observational environment at these locations may be poor. In other words, there are fewer satellites above the horizon than in other locations. Indeed, as shown in Figure 5.9, there are merely 8 - 11 visible satellites over the continental United States with elevation above 0° as compared to other

locations on the globe. Furthermore, most of the visible satellites are low elevation, resulting in a lower received carrier-to-noise ratio, rendering them more susceptible to loss of lock during the burst.

5.2.2 International GNSS Service Response

The International GNSS Service (IGS) is a voluntary federation of more than 200 worldwide agencies that submit GPS and GLONASS data for free, unrestricted use (*Dow et al.*, 2005). IGS also provides precise GPS and GLONASS products including precise satellite orbit determination and clock corrections. The data are primarily intended for the support of Earth science research, multidisciplinary actions, and education.

Each IGS station is directed under a set of guidelines that specify, among other things, that each receiver be capable of dual-frequency phase and code observations at a rate of at least one observation every thirty seconds. Furthermore, there are specifications and requirements for the hardware including antennas and radomes. Finally, each station must be permanent and continuously operating. The result is a highly reliable network that as of September 2007 has 345 operating stations.

The IGS network provides a view of the impact of the solar radio burst on a global perspective. By counting the number of simultaneous L1 and L2 code observations to each visible satellite that each receiver made at a particular point in time, a measure of the receiver's performance during the solar radio burst can be made. Recall that a receiver cannot produce a navigation solution when it is tracking less than 4 satellites. As shown in Figure 5.10, at the peak intensity of the solar radio burst on 6 December 2006, over half the IGS receivers on the sunlit

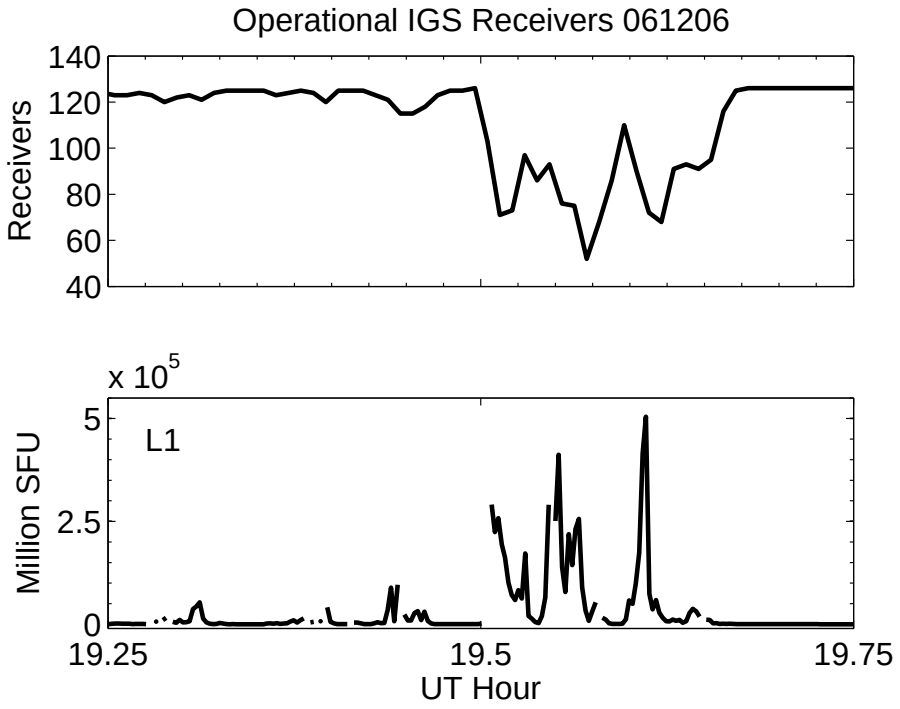

Figure 5.10: Number of operational sunlit IGS receivers on 6 December 2006 during the solar radio burst. Before and after the solar radio burst there were over 120 operational receivers. At the peak intensity of the solar radio burst, less than 60 receivers were producing dual-frequency code solutions.

hemisphere of the Earth were impacted to the point that they no longer produced a dual-frequency code navigation solution.

Each panel of Figure 5.11 illustrates the response of a random sampling of receivers in the IGS network. A movie (IGS.mov) is included in the accompanying CD that shows a time-lapse history of the observations at the same stations. Each dot represents the location of an IGS receiver and the color represents the number of dual-frequency code observations at that station. Across the bottom, the OVSA RHCP solar radio burst flux density at 1.6 GHz is plotted. Notice that, in general, the receivers that are closest to the subsolar point have fewer observations, indicating that they are more impacted by the solar radio burst. Again, this result is expected because the antenna gain is higher at high elevation angles, so the solar

radio burst has more impact on the receiver.

Most IGS users rely on the dual-frequency measurements produced by the receivers. As was indicated earlier, semi-codeless dual-frequency receivers are more prone to solar radio bursts because of squaring loss and the cross-correlation technique used to track the encrypted military signal on L2. How would a standard civilian (L1 only) receiver operate? Repeating the analysis above, but for L1 only, the panels in Figure 5.12 depict the response of the IGS receivers at L1 alone. The panels are at the same time and on the same scale as in Figure 5.11. The panels are also available as a movie (IGS_L1.mov) on the accompanying CD.

L1-only observations are not nearly as impacted by the solar radio burst as the dual-frequency observations. Since the L1 C/A code is a publicly available signal, the code sequence is known and is tracked directly. Thus, the cross-correlation and squaring losses are not as significant and will not affect the receiver's ability to track the signal as severely. Furthermore, the L1 C/A code is transmitted 3 dB higher than the P code on L2 and 6 dB higher than the P code on L1[2].

It is likely that most commercial, low-cost GPS receivers such as those used in vehicle navigation systems, hand-held devices including mobile phones, and other non-differential L1-only receivers worse performance than seen on the L1-only observations from the IGS network. The performance is likely worse because low-cost receivers do not employ mitigation techniques for interference and weak signal environments. It would be difficult to test the performance of most of these receivers because the GPS observables (pseudorange, carrier phase, carrier-to-noise ratio) are not user-accessible.

[2]The vehicles that are currently being launched, known as Block IIR-M and the next generation Block IIF will transmit the P code on L1 and L2 at the same power, only 3 dB lower than the L1 C/A code.

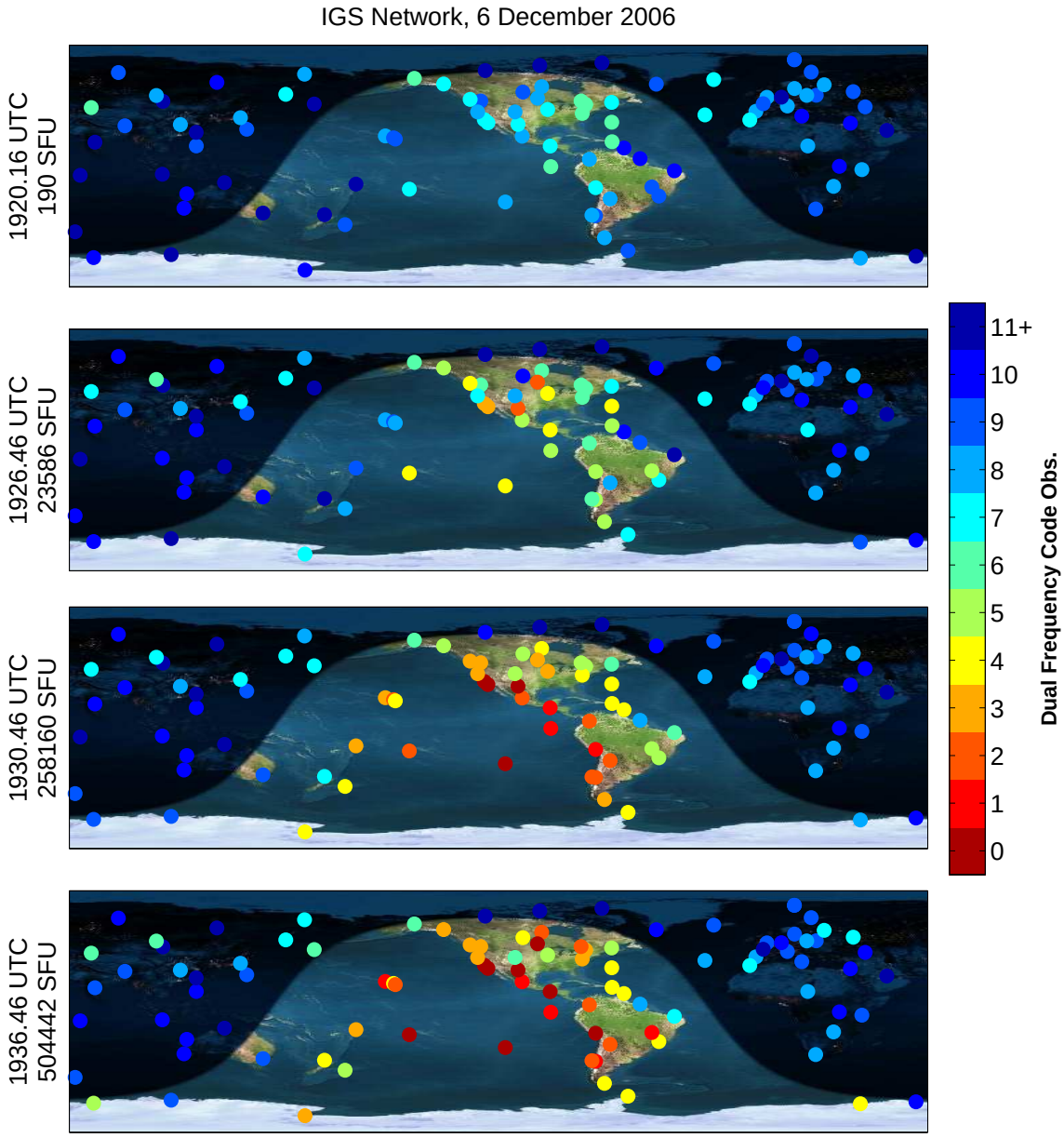

Figure 5.11: Each panel illustrates the number of dual-frequency code observations a random sampling of IGS receivers made at particular times (denoted on the x-axis). In addition to the time, the solar radio burst power at GPS L1 is indicated. At the peak of the burst, at 1936.46 UT, it is clear that the receivers most impacted were nearer to the subsolar point. Image of Earth © The Living Earth, 1996 and is used here by permission of the publisher. Day/night overlay created using Earth Viewer by John Walker.

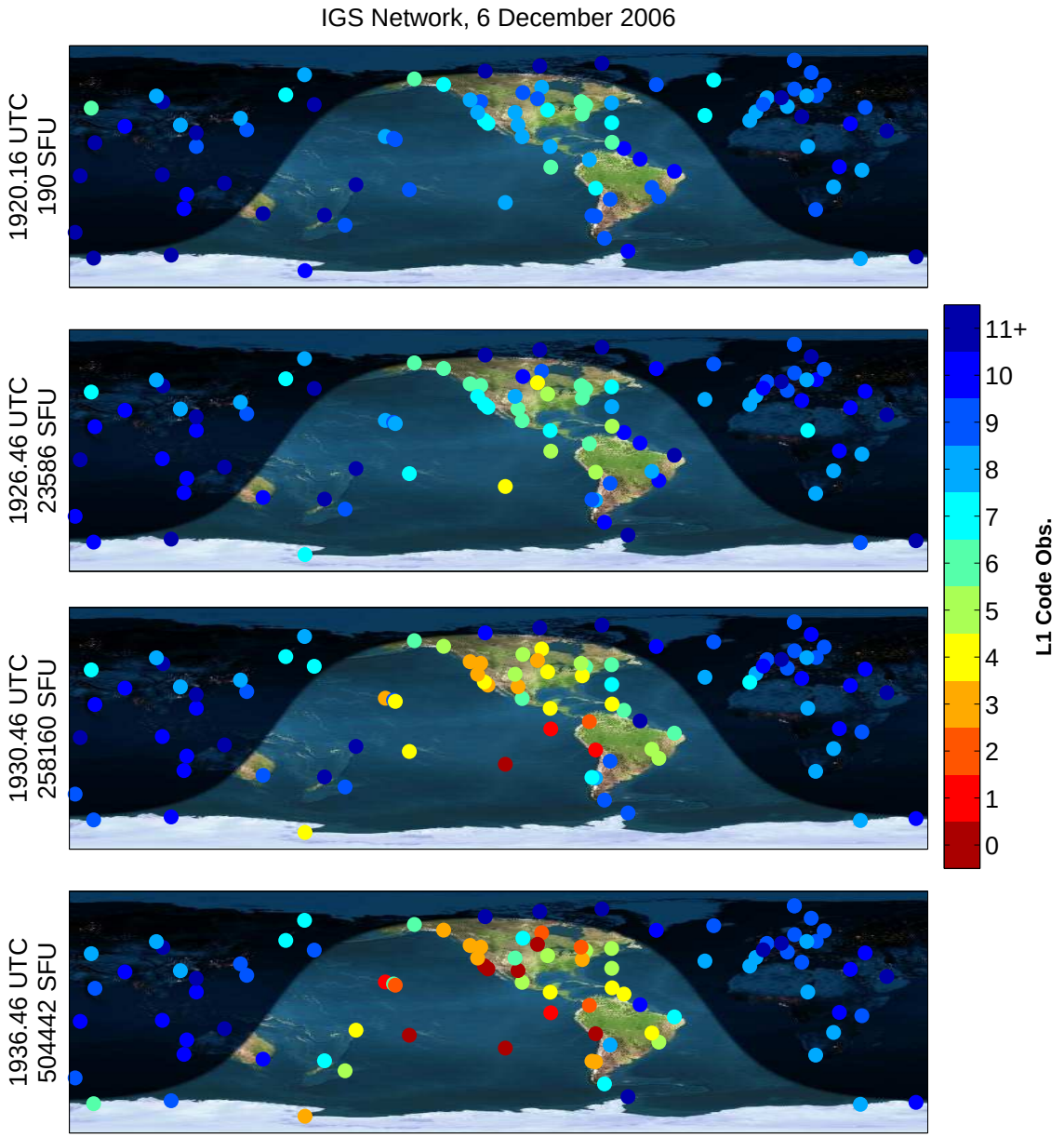

Figure 5.12: Each panel illustrates the number of L1 C/A code observations a random sampling of sunlit IGS receivers made at particular times (denoted on the x-axis). In addition to the time, the solar radio burst power at GPS L1 is indicated. L1-only observations were not as nearly impacted by the solar radio burst primarily for two reasons: the L1 C/A code is a public access code and is easily tracked by receivers, and the L1 C/A code is transmitted 6 dB higher than the P code on L2. Image of Earth © The Living Earth, 1996 and is used here by permission of the publisher. Day/night overlay created using Earth Viewer by John Walker.

5.2.3 Cornell ScintMon Receiver Response

Data typically used for scintillation studies is available from the Cornell ScintMon receivers operating at Arecibo Observatory in Puerto Rico. The data are worth mentioning because it is, to the author's knowledge, the first example of a solar radio burst causing a notable S_4 index. The S_4 index is a parameter used to estimate the intensity of amplitude scintillation typically associated with plasma irregularities. The S_4 index is defined as (*Yeh and Liu*, 1982):

$$S_4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}} \tag{5.1}$$

where I is the intensity of the signal. The term under the radical is merely the standard deviation of the signal amplitude divided by the mean signal amplitude. Typically, for GPS, the intensity, or carrier-to-noise ratio, is measured at 50 samples per second, and the S_4 index is averaged over a one-minute period.

In Figure 5.13, the carrier-to-noise ratio from PRN 28 on 6 December 2006 is shown as the gray line. The carrier-to-noise ratio for the previous day, shifted by a sidereal, day, is shown as the black line. There are several periods, denoted by arrows, where the receiver lost lock on the signal due to the solar radio burst. The Cornell ScintMon receiver typically loses lock for $C/N_0 < 27 - 29$ dB-Hz. Loss-of-lock, however, is also a function of the duration of reduced carrier-to-noise ratio, so rapid variations may not necessarily lead to complete loss of tracking.

Figure 5.14 illustrates the S_4 index as calculated from all satellites at Arecibo Observatory that were visible throughout the two-hour period from 18:00 - 20:00 UT. There is a strong similarity in the shape of the S_4 curves. Since the solar radio burst causes a similar carrier-to-noise ratio fade on all visible satellites, this result

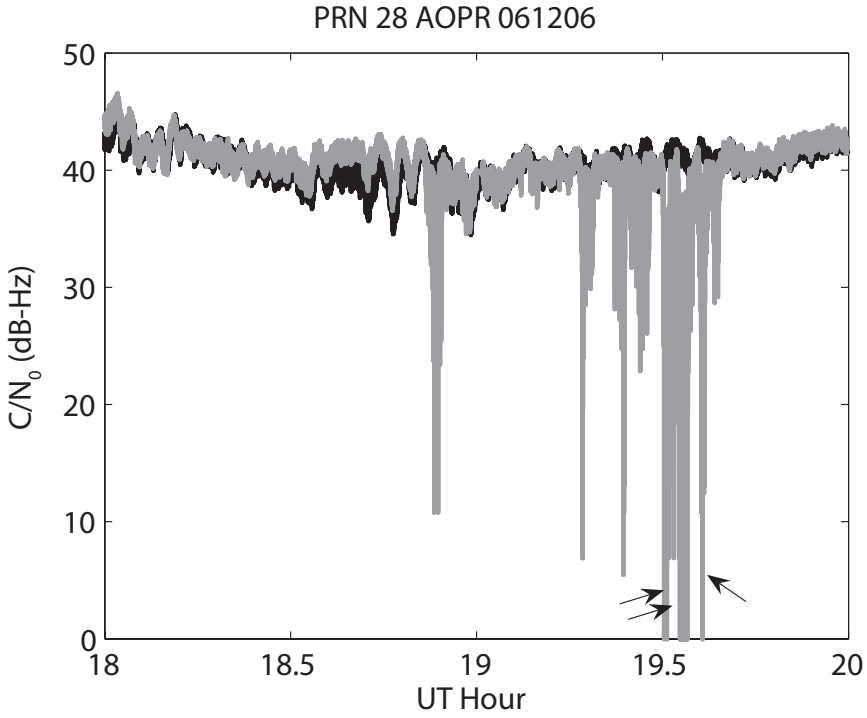

Figure 5.13: Carrier-to-noise measurements as recorded by the Cornell University ScintMon Receiver operating at Arecibo Observatory, PR for PRN 28. The gray line represents the amplitude data recorded on 6 December 2006, and the black line represents the data recorded on 7 December 2006. Note that the receiver went through several periods of losing lock, denoted by arrows.

is expected. It is important to note that the S_4 index is not a result of ionospheric scintillation. Ionospheric scintillation can actually occur during the day as a result of plasma irregularities associated with Storm Enhanced Density (SED) plumes (e.g. *Basu et al.* (2005b) and *Basu et al.* (2005a)), which are typically associated with geomagnetic activity, which is ultimately associated with solar activity. Thus, day-time C/N_0 fades should be analyzed from the perspective that they could be caused by solar radio bursts or irregularities. To reiterate, fades induced by scintillation differ from solar radio bursts because solar radio bursts will only cause decreases in C/N_0 from the nominally received C/N_0. Conversely, scintillations will cause both increases and decreases in the received C/N_0 as compared to the nominally received C/N_0.

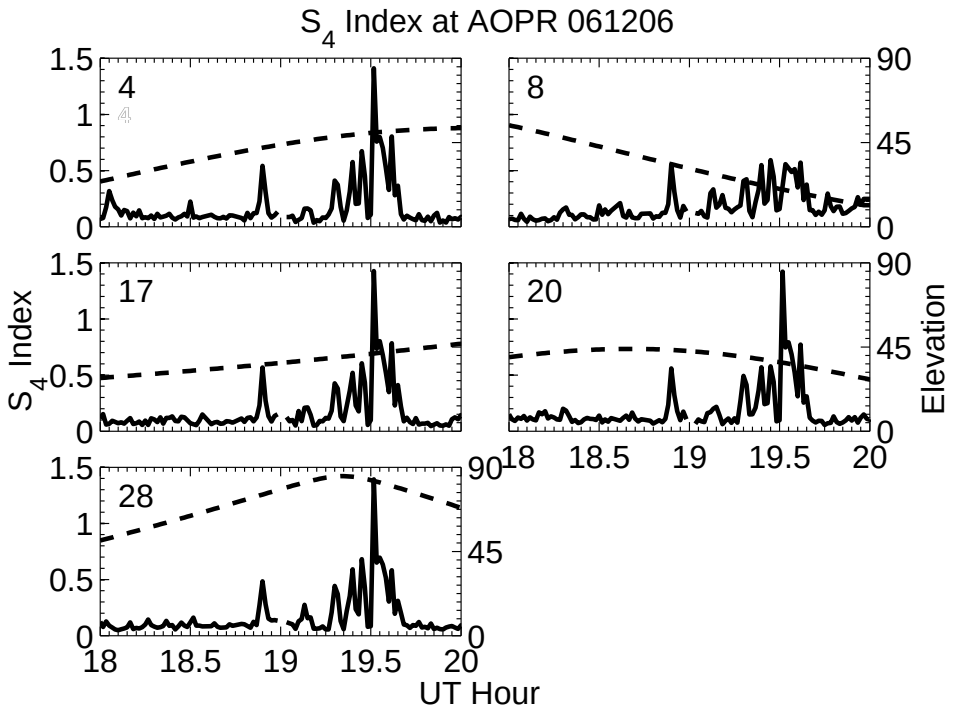

Figure 5.14: S_4 index for five GPS satellites as observed from Arecibo Observatory, PR for PRN 28. Notice the the strong similarity in the S_4 lines for each satellite.

5.3 WAAS Receiver Tracking Performance During the 6 December 2006 Solar Radio Burst

Up to this point, only the effects of the solar radio burst on recorded carrier-to-noise ratio have been explored. It has been clearly demonstrated that the fades can become so large that the receiver altogether loses lock on the satellite signal, especially for semi-codeless L2 receivers. What effect, however, does dropping a satellite have on the navigation solution? More importantly, what effect does the decreased carrier-to-noise ratio have on the quality of the pseudorange measurements and the resulting navigation solution? The Delay Lock Loop (DLL) is responsible for maintaining the code-phase of the PRN sequence properly aligned, and ultimately the pseudorange measurements are made from the DLL. As shown

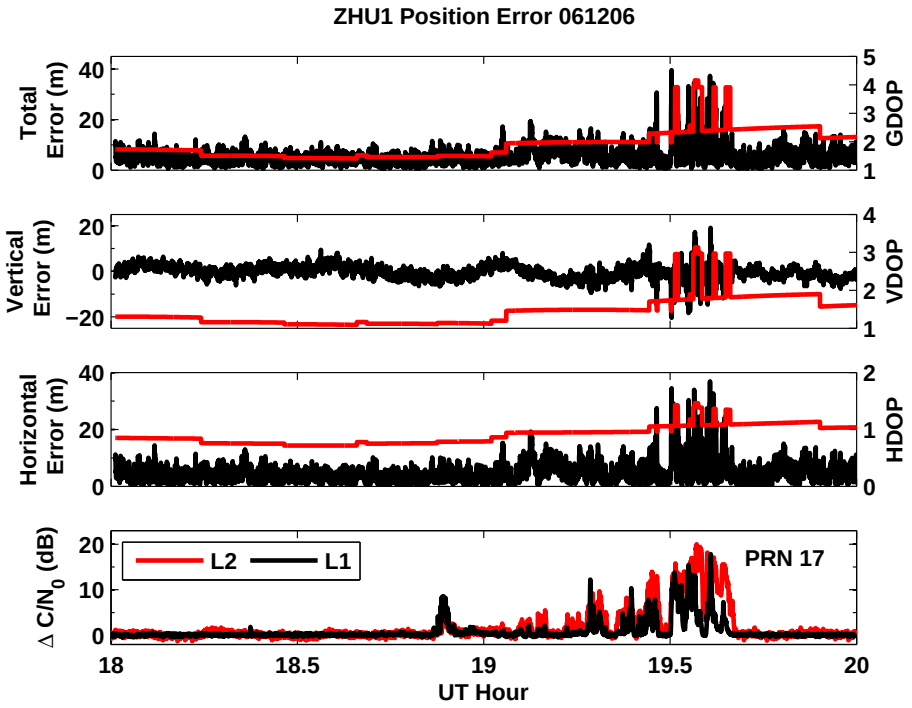

Figure 5.15: Position error (black line) as calculated from the Houston WAAS Reference Station (ZHU1). The red line depicts the relevant Dilution of Precision (DOP) parameter for each error plot (GDOP - Geometrical DOP, VDOP - Vertical DOP, HDOP - Horizontal DOP). In the bottom panel, the fades as observed on PRN 17 are plotted. A specialized dual-frequency receiver was used to minimize errors resulting from ionospheric propagation, tropospheric propagation, and satellite and ephemeris clock errors.

in Section 4.2.3, the code tracking error of the DLL is a function of the carrier-to-noise ratio. At decreased carrier-to-noise ratios it is expected that the receiver will exhibit greater positioning errors due to the increase in pseudorange errors.

The navigation error is calculated using all satellites above 5° elevation (to minimize multipath errors) for which dual-frequency code observations are available. The navigation solution is calculated with the standard dual-frequency ionosphere-free pseudorange as specified by the *IS-GPS-200D* (2004). The dual-frequency ionosphere pseudorange eliminates the bulk of the ionospheric delay at the cost of a slightly higher variance in noise. This technique, however, greatly reduces the

ionospheric error. Each pseudorange is subsequently corrected for tropospheric propagation errors with a Saastamoinen model and an appropriate obliquity factor. All relativistic corrections as specified by the *IS-GPS-200D* (2004) are also applied. Finally, the NASA/JPL final orbital ephemeris and clock parameters minimize errors originating from satellite clock drift and ephemeris inaccuracy. These parameters reduce satellite positioning error to less than 5 cm and satellite clock errors to less than 0.1 ns. The navigation solution and Dilution of Precision (DOP) is then calculated as detailed in Chapter 3, Section 3.7.5. The resulting errors are highly indicative of the receiver performance since all other error sources are minimized.

Figure 5.15 illustrates the marked increase in positioning errors during the strongest portions of the solar radio burst on 6 December 2006 at the Houston (ZHU1) WAAS Reference Station. A movie entitled 'Error.mov' is found on the accompanying CD and illustrates the horizontal (left panel) and vertical (middle panel) position error as well as the satellite constellation (right panel) as a function of time for all satellites above 5°. Across the bottom, the carrier-to-noise ratio fades as observed at L1 (black) and L2 (blue) are illustrated. Finally, on the right, the time, number of visible satellites, and the HDOP and VDOP are displayed.

In the top panel of 5.15, the overall position error is indicated as the black line, with the Geometrical DOP as the red line. Prior to the onset of the solar radio bursts at 1848 UT, the mean position error is 4.5 m with a standard deviation of 1.9 m. The errors then climb drastically during the peak periods of the solar radio burst. From 1924 UT to 1948 UT, the mean error climbs to 7.4 m, with a peak error of 39.5 m at 1929 UT. At 1933 UT, a minimum of 5 satellites are tracked and included in the navigation solution for a period of thirty seconds. At

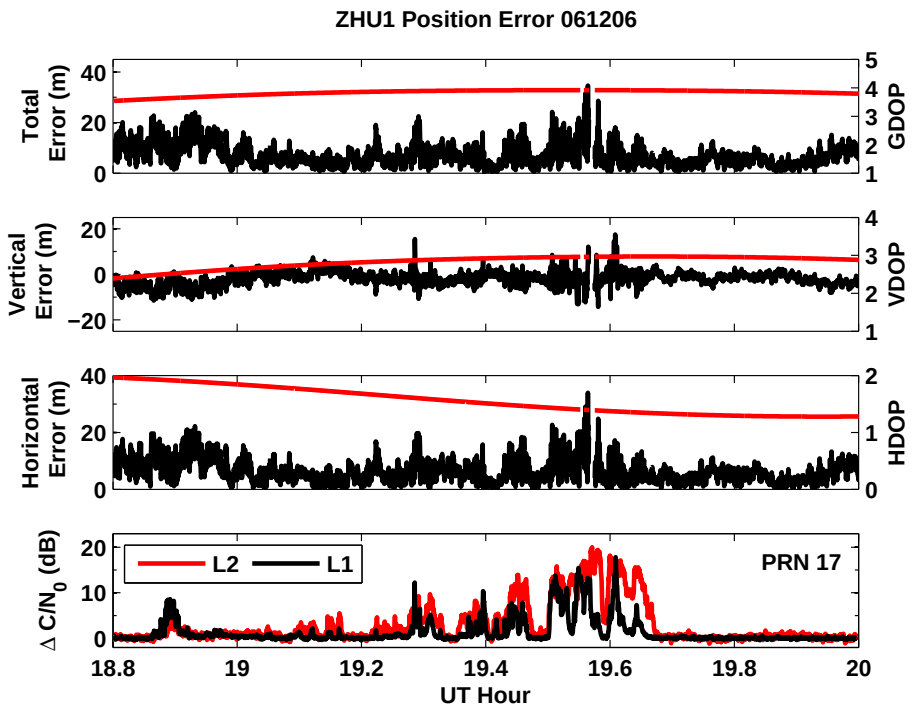

Figure 5.16: Position error (black line) as calculated from the Houston WAAS Reference Station (ZHU1) using only PRN 2, 4, 9, 17, 24, and 28. These six satellites were tracked throughout the burst, with the exception of a short period of time around 1933 UT. The red line depicts the relevant Dilution of Precision (DOP) parameter for each error plot (GDOP - Geometrical DOP, VDOP - Vertical DOP, HDOP - Horizontal DOP). In the bottom panel, the fades as observed on PRN 17 are plotted. The DOP was relatively constant throughout the solar radio burst, however, the position errors still increased substantially at the peak fades indicating that the DLL suffered from increased tracking errors.

all other times, the receiver was tracking at least 6 satellites. The middle panels of Figure 5.15 illustrate the local vertical and horizontal position errors, along with the Vertical DOP and the Horizontal DOP. Note in particular that the VDOP parameter increases substantially as low elevation satellites are dropped during the solar radio burst. Low elevation satellites are particularly prone to being dropped because the received carrier-to-noise ratio from these satellites is already substantially lower than high elevation satellites. Recall from Chapter 3, Section 3.7.5 that the variance in positioning error is proportional to DOP.

At the peak of the solar radio burst, the Houston WAAS Reference Station was tracking PRNs 2, 4, 9, 17, 24, and 28. The position error of the receiver was calculated again with only these satellites in an effort to eliminate reduced DOP as a culprit of increased navigation errors. The results of this calculation are shown in Figure 5.16.

The mean position error during quiet periods is 5.62 m with a standard deviation of 2.86 m. The mean position error has climbed slightly because 6 satellites are included in the navigation solution, and a 6-satellite DOP is always larger than a DOP for which more than 6 satellites are available. The navigation errors climb substantially during the solar radio burst even while the DOP remains relatively constant. From 1924 UT to 1948 UT, the mean error climbs to 6.32 m (standard deviation of 4.35 m), with a peak error of 34.59 m at 1933 UT. For all the plots, the DOP remains relatively constant, but there is clearly an increase in the error corresponding to the peak fades on L1 and L2, corresponding to an increase in the tracking error variance of the DLL.

Figure 5.17 illustrates the residual pseudorange error for PRN 17 (top panel) and PRN 24 (middle panel) respectively. PRN 17 is the satellite with the highest carrier-to-noise ratio during the solar radio burst and PRN 24 is the satellite with the lowest overall carrier-to-noise ratio that does not lose lock during the solar radio burst. Residual pseudorange error is the measurement error left in the pseudorange after the exact satellite position has been computed, satellite and receiver clock biases have been accounted for, and all propagation errors (ionospheric, tropospheric, relativistic) have been accounted for. The increase in residual errors at the time of the deepest fades is direct proof that the DLL is suffering from increased code tracking error.

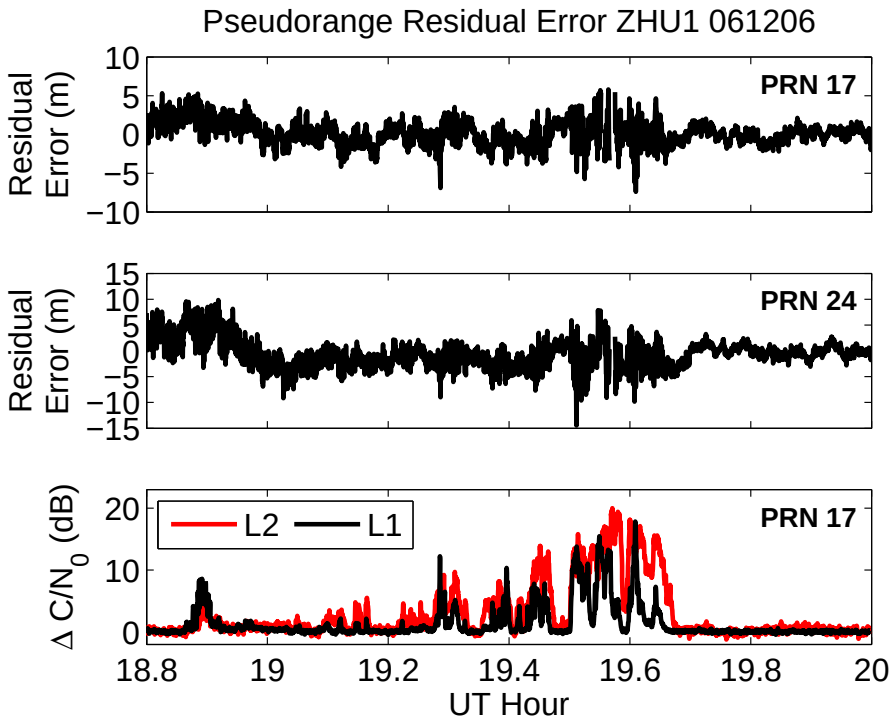

Figure 5.17: Residual pseudorange error for PRN 17 (top panel) and PRN 24 (bottom panel). In the bottom panel, the fades as observed on PRN 17 are plotted. The residual pseudorange error climbs substantially during the peak of the solar radio burst, indicating that the DLL suffered from degraded tracking performance. The large residual error prior to 19 UT for PRN 24 is a result of multipath due to the low elevation angle of the satellite.

For this analysis, only the pseudorange (code) navigation solution is considered. Fortunately, the carrier phase observations from each satellite can be combined with the pseudoranges using a recursive filter to produce a carrier-aided navigation solution. The carrier-phase is substantially less affected by multipath and also suffers from less measurement variance at low carrier-to-noise ratios. The carrier-aided navigation solution will substantially lower the variance in the navigation solution error.

5.4 Other December 2006 Solar Radio Bursts

Active region 10930 produced two more X-class solar flares and subsequent solar radio bursts that are worth mentioning because each of these solar radio bursts, although not as large as the 6 December 2006 event, exceeded the previous known records at the GPS frequencies.

The first event was on 13 December 2006 at approximately 0240 UT when active region 10930 unleashed an X3 solar flare. Again a solar radio burst accompanied the solar flare, and there was a corresponding decrease in the received carrier-to-noise ratio from the satellites. Figure 5.18 demonstrates the fade in carrier-to-noise ratio for the event as measured on the FAA WAAS receiver located in Honolulu, Hawaii (21.3°N, 157.9°W).

Data from the OVSA is not available for this event because the array was on the night-side of the earth. Instead, data from the Nobeyama Radio Observatory in Japan are available. In Figure 5.18, the RHCP solar radio burst power at 1.0 and 2.0 GHz is shown as the red and black line respectively. The peak solar radio burst power at 1.0 GHz is 248,000 SFU at 0228.08 UT. At 2.0 GHz, the peak solar radio burst power is 276,000 SFU at 0335.52 UT. The peak fades observed on GPS are 4 dB on the L1 frequency at 0228.43 UT, and 7.9 dB on the L2 frequency at 0230.57 UT. The 4 dB fade at L1 corresponds to a 55,000 SFU RHCP solar radio burst, and the 7.9 dB fade at L2 corresponds to a 184,000 SFU RHCP solar radio burst. Each of the fades associated with the solar radio burst had a duration less than 5 minutes. The large discrepancy in the derived solar radio burst power from the GPS data and the NRO data are most likely due to the large difference in the measurement frequencies from the GPS frequencies.

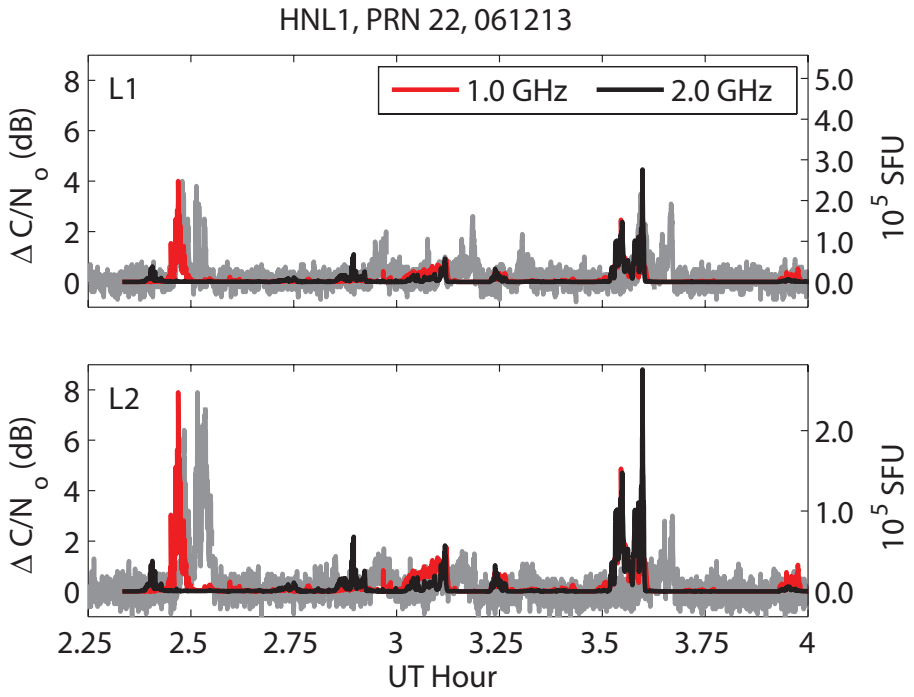

Figure 5.18: 13 December 2006 solar radio burst and GPS response as measured on PRN 22 by the FAA WAAS receiver in Honolulu, Hawaii (21.3°N, 157.9°W). The Owens Valley Solar Array was not illuminated by the sun. Instead, RHCP solar radio burst measurements from the Nobeyama Radio Observatory in Japan are shown. The array makes measurements at 1.0 and 2.0 GHz as shown by the red and black lines, respectively. The peak fade at L1 is approximately 4 dB at 0228.43 UT and the peak fade at L2 s approximately 7.9 dB at 0230.57 UT. At 0230 UT the sun was at an elevation of about 16°.

The last known solar radio burst associated with active region 10930 accompanied the X1 solar flare on 14 December 2006 at 2215 UT. The accompanying solar radio burst is depicted in Figure 5.19. At 1.6 GHz, the peak RHCP solar radio burst intensity is 107,000 SFU at 2252.40 UT, whereas at 1.2 GHz, the peak intensity is merely 2,000 SFU at 1046.35 UT. Again, this solar radio burst is particularly noteworthy because, even at 107,000 solar flux units RHCP, it is still larger than most other solar radio bursts previously recorded at these frequencies.

In Figure 5.20, the response of the FAA WAAS receiver located in Honolulu, Hawaii is shown. The peak fade at L1 is 10 dB at 2304.32 UT, corresponding to a

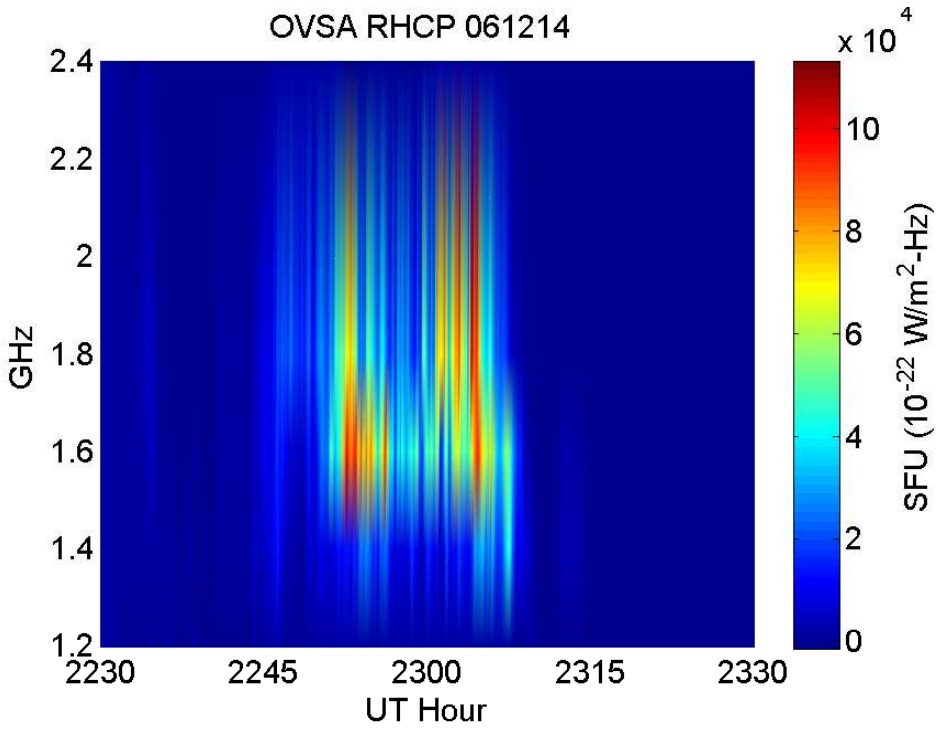

Figure 5.19: 14 December 2006 solar radio burst as recorded at the Owens Valley Solar Array. The data shown are for the RHCP mode only. The burst was particularly energetic at the GPS L1 frequency.

RHCP solar radio burst intensity of 75,000 solar flux units. The fade at L1 had a duration longer than 15 minutes. The fade at L2 is negligible and the variations seen in the plot are due to a changing multipath environment over successive days. The data, however, illustrate that solar radio bursts can be extremely dynamic in both their duration and intensity over short time and frequency scales. This dynamic nature of the solar radio burst can be exploited to aide the receiver tracking loops and will be further discussed in the next chapter.

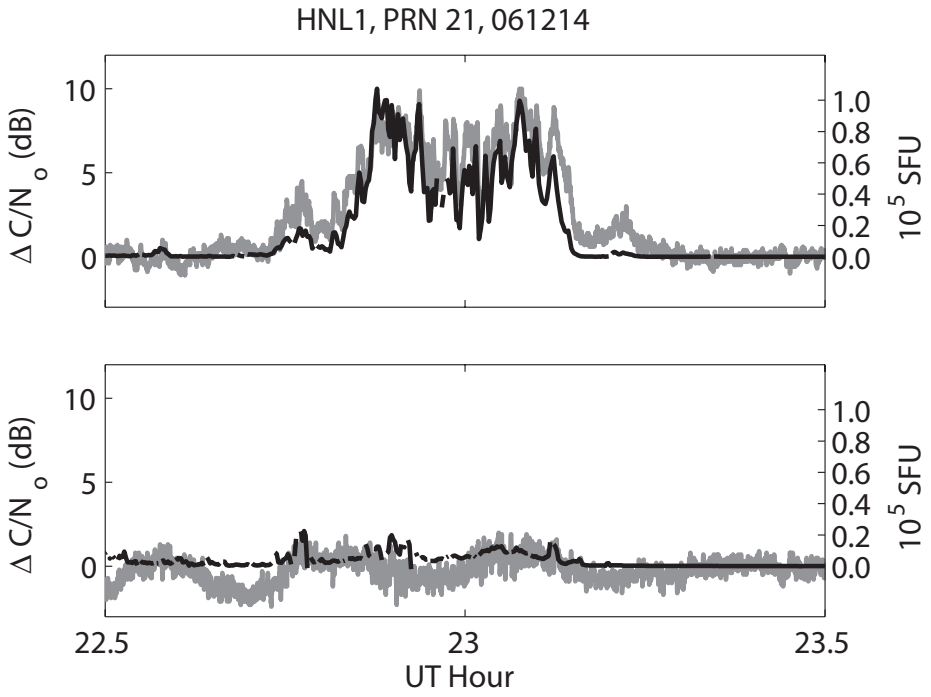

Figure 5.20: Fades induced by the 14 December 2006 solar radio burst as measured on PRN 21 from the FAA WAAS receiver located in Honolulu, Hawaii. The fade at L1 is 10 dB and had a duration of longer than 15 minutes. The fade at L2, on the other hand, is negligible and is buried with the multipath effects. The sun was at an elevation of 44.5°.

5.5 Conclusion

The chapter presents a study of the record-setting solar radio bursts of December 2006. All three of these solar radio bursts were larger than any solar radio burst previously recorded at these frequencies during solar minimum. The 6 December 2006 solar radio burst is the largest burst on record, with independent flux estimates of 500,000-2,000,000 solar flux units in the right-hand circularly polarized mode at GPS L1.

The solar radio burst caused fades of greater than 15 dB at the L1 frequency on the particularly robust Federal Aviation Administration's (FAA) Wide Area Aug-

mentation System (WAAS) network of reference receivers, and the loss of observations from several satellites briefly impacted the precision approach service. Other networks of receivers, including the International GNSS Service (IGS) network and the Cornell ScintMon network suffered total loss of lock on many satellites, and in many instances total loss of navigation solution for a brief period of time.

All sunlit receivers were significantly impacted and positioning errors climbed dramatically both due to and increase in the Dilution of Precision when satellites are dropped due to weak carrier-to-noise ratios, as well the decreased received signal power from the satellites, which causes an increase in code tracking errors. The December events prove that solar radio bursts can have a significant impact on GNSS receivers, and mitigation techniques should be considered for robust, uninterrupted operation.

Acknowledgements

Dale Gary of the New Jersey Institute of Technology provided both the OVSA and FST data for the December 2006 solar radio bursts and results of the detailed analysis of the 6 December 2006 event. Gratitude is also extended to Pat Reddan of Zeta Associates Inc. for providing valuable resources on WAAS, including antenna gain patterns and references to papers. 1-second WAAS data was provided by Tom McHugh of the FAA and also by Patricia Doherty of Boston College. The views reported here are solely the views of the author and not of the FAA or other government agencies.

Chapter 6

Mitigation and Future Work

As for the future, your task is not to foresee it, but to enable it. - Antoine de Saint-Exupery

6.1 Introduction

A study conducted by *Csillaghy and Benz* (1993) studied 196 individual millisecond duration solar radio bursts and measured their full-width half maximum (FWHM) bandwidths. Three individual events (consisting of several bursts) were measured in the GPS frequency ranges, and the FWHM bandwidth was never measured to be narrower than 17 MHz. These numbers are considered to be rather narrow as compared to the general case (private communication with Dale Gary). The solar radio bursts of December 2006 were many small events of 3 - 4 MHz in bandwidth that in several cases overlapped to create even wider interference bandwidths. Radio interference with bandwidths greater than 2 MHz (20 MHz) are considered wide-band noise for the L1 C/A code (L1/L2 P(Y)). Radio interference with bandwidth less than 2 MHz are typically considered narrow-band interferers and have effects that uniquely affect the L1 C/A code due to spectral line interference (*Spilker and Natali*, 1996). Solar radio bursts, in general, are a wideband interference source.

This chapter will introduce several techniques to mitigate broad-band interference lying within the GPS frequency bandwidth. All of these techniques require extensive modifications to existing receiver hardware, or more commonly,

*Portions of this chapter are based on the publication *Cerruti et al.* (2007), which is joint collaborative work with A. P. Cerruti, P. M. Kintner, D. E. Gary, A. J. Mannucci, R. F. Meyer, P. Doherty, and A. J. Coster. Reproduced by permission of the American Geophysical Union.

are stand-alone solutions. The goal of this chapter is not to give a thorough, detailed explanation of each technique, but rather to give the reader a starting point from which the various techniques could be wielded to mitigate the effects of a solar radio burst.

Next, in Section 6.3, a technique is developed whereby solar radio burst flux is estimated from the fades measured at the WAAS Reference Stations. The WAAS Reference Stations are chosen over other networks of receivers (e.g. National Geodetic Survey's Continuously Operating Reference Stations (CORS) or the International GNSS Service (IGS)) because the WAAS Reference Stations all employ identical hardware, thus facilitating analysis. Sadly, there does not exist a standard that receiver manufacturers abide by to consistently compute, record, and report carrier-to-noise ratios. Unlike the WAAS network, the CORS and IGS networks are a hodgepodge collection of receivers that, while they all conform to strict standards, are of different manufacturers and different models. Hence, the receivers utilize different tracking loop bandwidths, antennas (and hence gain patterns), automatic gain control (AGC), and algorithms that can influence carrier-to-noise ratio measurements. Fortunately, the antenna gain pattern has the largest effect on the carrier-to-noise measurements, where as the tracking loop bandwidth and the AGC typically only affect rapid transients in carrier-to-noise measurements. For consistency and as an example, the technique to obtain the solar radio flux from GPS carrier-to-noise ratios is done from the WAAS receivers, but it should be a straight-forward extension to include measurements from other receivers.

The technique of obtaining solar radio flux measurements from GPS receivers is pertinent for a variety of reasons. First and foremost, GPS receivers are distributed globally, offering continuous, redundant measurements that can be verified

and compared with other instruments. Secondly, the dynamic range and fast response of a properly operating GPS receiver makes them particularly suited to measuring solar radio flux intensity, which as was shown in the December 2006 events, can vary dramatically. These measurements can be compared to and augment measurements from dedicated observatories such as the Owens Valley Solar Array (OVSA), the Nobeyama Radio Observatory (NRO), and the United States Air Force (USAF) Radio Solar Telescope Network (RSTN). The solar radio flux measurements obtained from WAAS will be shown to agree well with the OVSA data. Finally, GPS measurements could be used to verify the accuracy of the solar radio burst historical record.

The last portion of this chapter, Section 6.4, briefly discusses the context of the December 2006 solar radio bursts with respect to the historical record. The strength of the December 2006 solar radio bursts came as a surprise to many scientists because they came at solar minimum and easily broke all other records at these frequencies. As shown by *Nita et al.* (2002), solar radio bursts typically follow a power-law distribution of intensity versus occurrence, and the December events suggest that other solar radio bursts of their magnitude could have been observed in the 40-year historical record. *Nita et al.* (2002) suggests that the larger events may not have been recorded because of equipment saturation or perhaps because of some limiting process in the generation of solar radio bursts. Many instruments, including the OVSA and particularly the USAF RSTN saturated during the December 2006 solar radio bursts, suggesting that the historical record is inaccurate, and could be, in part, amended using GPS measurements and the technique suggested in Section 6.3.

6.2 Mitigation

A naive approach to mitigate the effects of a solar radio burst would be to attempt to modify the antenna gain pattern or reducing the system noise temperature since they are the main parameters that vary in Equation 4.6. These attempts are generally not practical.

Decreasing the antenna gain would reduce the solar radio burst impact but also results in decreased received signal power from the satellite. Furthermore, an advantage of high-gain antennas is that they have inherently better multipath rejection, resulting in better ranging precision and lower navigation solution errors while guaranteeing complete coverage of the sky.

Lowering the noise temperature of the receiver is always beneficial to carrier-to-noise ratio. However, fancy techniques such as cooling the receiver and amplifiers with liquid nitrogen to reduce the system noise temperature of the receiver are impractical as most receivers are destined for low-cost, low-power applications such as cell phones and hand-held units. An additional undesired effect of reducing the noise figure is that the receiver actually becomes more sensitive to solar radio bursts. Other more practical methods of mitigating the effect of the solar radio burst on the receiver need to be considered.

At the very least, a receiver should have some method to warn the user of the presence of interfering signals. In many receivers, carrier-to-noise ratio is not available or is difficult to interpret. *Ward* (1996b) emphasizes that a GPS receiver that has been enhanced to operate in the presence of RF interference has a built-in RF interference detector. An RF interference meter simply requires that the AGC control voltage levels at the final IF output must be digitized and provided to

the receiver to be interpreted. Determining the presence of interference is simple because the control voltage level in the presence of interference will be significantly different than that of the unaffected control voltage level.

Several additional ways of mitigating the effects of interference are possible. The first line of defense against solar radio bursts is through careful design of the tracking loops to have as narrow a bandwidth as possible. Narrowing the noise bandwidth of the phase lock loop can significantly improve the receiver's immunity to wide-band interference, but at the same time can make the receiver more vulnerable to satellite-receiver dynamics (e.g. acceleration or jerk). Improvements to the receiver clock can also improve the tracking error, thus reducing susceptibility to low carrier-to-noise ratios. Finally, there exist carrier tracking loops known as frequency lock loops that have a substantial advantage over phase lock loops in the presence of interference. These techniques are described in Section 6.2.1.

Bit prediction is another technique whereby the susceptibility of the tracking loops to low carrier-to-noise ratios can be significantly improved. Discussed in Section 6.2.2, bit prediction allows the tracking loops to operate such that squaring loss is removed from the tracking error. Furthermore, bit prediction allows for longer integration times in low-dynamics applications, further decreasing the minimal carrier-to-noise threshold.

Another line of defense is to include navigation aids from outside instruments such as Inertial Measurement Units (IMU), radar, altimeters, and compasses. Most of these units are immune to radio-frequency interference, perhaps with the exception of radar. These enhancements are discussed in 6.2.3.

Next, more drastic measures of modifying receiver hardware are discussed in

Section 6.2.4. The choice of pre-correlation filter (which determines the signal bandwidth within the receiver) and number of digitization levels can have small, but appreciable effects on the overall implementation losses within the receiver.

Perhaps the best way to mitigate interference is to use a nulling antenna array, which can electronically steer a null in the beam-pattern of the antenna towards an interfering source. Significant improvements in overall carrier-to-noise ratio can be achieved with nulling antenna arrays, but at a cost of increased receiver complexity. Nulling antenna arrays are discussed in Section 6.2.5.

Finally, several improvements can be made to the satellites via the addition of new frequencies, increasing transmitter power, or the addition of new signals on existing frequencies (Section 6.2.6). These improvements are the most costly, requiring not only extensive receiver modifications, but also replacement of satellites in orbit. Fortunately, the next generation of GPS satellites are slated to have new signals, and higher transmission power. Furthermore, the push to reinstate GLONASS as a functional system, as well as the addition of the European Galileo and Chinese Beidou systems will greatly increase the number of visible satellites, thus decreasing the susceptibility of the receiver to stop navigating during moderate solar radio bursts.

Common techniques for mitigating narrow-band interference such as adaptive filter arrays[1] are ineffective against wide-band interference and will not be discussed. Furthermore, solar radio bursts are relatively long-duration events, and so pulse-blanking techniques, although they may be useful during some solar radio burst events, are generally impractical[2].

[1]Adaptive filters are digital filters with multiple taps, where the weight of the tap can be adjusted using an algorithm to efficiently filter out narrow-band interferers from the digitized GPS signal.

[2]Pulse-blanking is a technique whereby the GPS signal within the receiver is turned off for

6.2.1 Carrier Tracking Loop Improvements

First introduced in Chapter 4, Section 4.2.3, the phase lock loop (PLL) and delay lock loop (DLL) work in tandem to track the carrier and code phase, respectively, of each satellite signal. The PLL loses lock at a much higher threshold (typically 25 - 30 dB-Hz for the L1 C/A code) than the DLL (typically 15 - 20 dB-Hz). Furthermore, carrier aiding from the PLL eliminates dynamic stress error in the DLL. As a result, if loss of lock on the PLL occurs, the DLL will lose lock soon thereafter. Thus, to minimize the effects of noise on GPS, the first line of defense is to improve the PLL performance by minimizing the tracking errors.

Recall that the 3-σ threshold for maintaining phase lock is about 45°, which corresponds to 15° at 1-σ. In Chapter 4, Section 4.2.3, only the phase error introduced by the tracking loop thermal noise was considered. In reality, the phase error is a function of several parameters and is computed as (*Ward*, 1996a):

$$\sigma_{PLL} = \sqrt{\sigma_{phi}^2 + \sigma_v^2 + \theta_A^2} + \frac{\theta_e}{3} \leq 15^\circ \tag{6.1}$$

where σ_{phi} is the thermal noise (explored in Chapter 4, Section 4.2.3), σ_v is the vibration- induced oscillator jitter, θ_A is the Allan variance-induced oscillator jitter in degrees due to the receiver clock, and θ_e is the dynamic stress error caused by receiver-satellite dynamics. Each parameter can be modified appropriately to improve the tracking performance and decrease the sensitivity of the tracking loop to outside noise influences. Based on *Ward* (1996a), each of the improvements will be described in the following subsections.

short durations to overcome intense, short-duration interference as associated, for example, when operating near a pulsed radar.

PLL Noise Bandwidth

Proper selection of the noise bandwidth of the PLL has the most significant effect when trying to minimize the effects of interference. Recall from Equation 4.8 that the overall phase error is given as:

$$\sigma_\phi = \frac{180}{\pi}\sqrt{\frac{B_L}{C/N_0}S_L} \text{ degrees} \tag{6.2}$$

and is a linear model of tracking error as a function of the squaring loss factor, S_L, the carrier-to-noise ratio C/N_0 of the incoming signal, and the single-sided noise loop bandwidth, B_L (Hz) (*Woo*, 2000). The current constellation only provides signals with data modulation present, and so, barring specialized techniques (some of which will be explored in Section 6.2.2, all GPS receivers suffer from squaring loss, S_L. Recall that the squaring loss, as given by Equation 4.9, is:

$$S_L \cong 1 + \frac{B_P}{2C/N_0} \tag{6.3}$$

where B_P represents the pre-correlation bandwidth of the signal, and C/N_0 is again the carrier-to-noise ratio.

A paradox must be solved by the GPS engineer for their particular operation. To tolerate dynamic stress error, the pre-correlation integration time $T = 1/B_P$ must be short, and the loop filter bandwidth, B_L, should be wide. On the other hand, to minimize tracking errors, the pre-correlation integration time should be large, and the loop filter bandwidth should be wide. An outside noise source, such as a solar radio burst, will substantially increase the noise in the phase lock loop, rendering carrier tracking more difficult. One defense is to decrease the PLL noise

loop bandwidth, rendering the PLL less sensitive to increases in noise but more vulnerable to dynamics.

For GPS applications, at least a second-order PLL is used to accurately track relative receiver-satellite velocity (Doppler shift). Second-order PLL's are adequate for stationary receivers, but for moving platforms, the second-order PLL has a constant error in the presence of receiver-satellite acceleration. A third-order PLL is ideally used in GPS applications because it can completely correct for receiver-satellite acceleration, regardless of the noise loop bandwidth, B_L. The third-order PLL is susceptible to receiver-satellite jerk (rate of change of acceleration), but for most applications the jerk is usually minimal (*Ward*, 1996a). The 1-σ 15° maximum phase error is exceeded at the same point in carrier-to-noise ratio for any order tracking loop when all other parameters are held equal.

When the loop order is increased, there is an improvement in dynamic stress performance, and the noise loop bandwidth can be reduced for a given application. When using pre-correlation integration times of 20 ms (the maximum permissible since this is the bit-rate of the GPS navigation message) and a bandwidth of 15 Hz, the 1-σ threshold of 15° is exceeded at 23.8 dB-Hz. Reducing the bandwidth to 3 Hz reduces the threshold to 17.9 dB-Hz, an improvement of almost 6 dB.

Of course, increasing the loop order comes at an additional cost of implementation complexity, and also stability. While second-order loops are stable at all bandwidths, third-order loops are prone to stability issues when the loop bandwidth exceeds 18 Hz. For most GPS applications, a loop bandwidth of 18 Hz, however, is much wider than necessary, and corresponds to maximum tolerable line-of-sight jerk dynamics of over 540,000 deg/sec^3 (equivalent to over 30 G/sec jerk) (*Ward*, 1996a). Jerk performance of the PLL is explored in the Appendix,

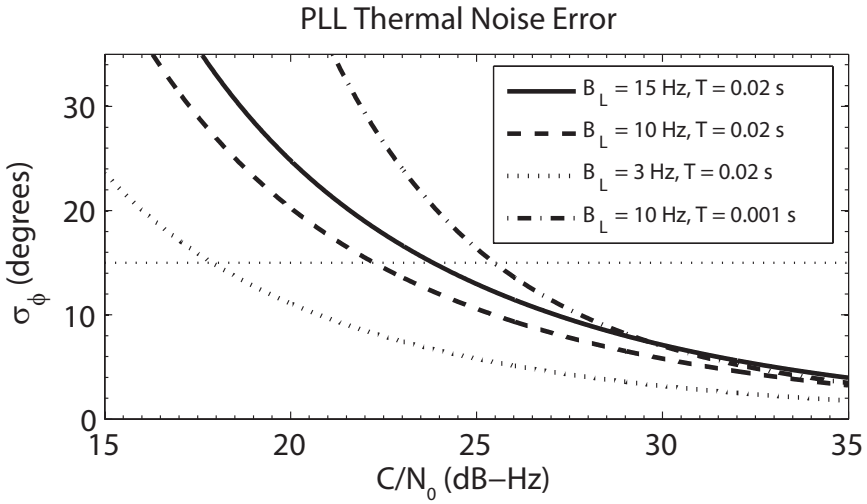

Figure 6.1: The phase error as a function of PLL loop bandwidth for a PLL. The lines are representative of any order PLL, but a second-order PLL is susceptible to receiver-satellite acceleration, and a third-order PLL is susceptible to receiver-satellite jerk. The horizontal dotted line represents the maximum tolerable error of 15°. Reducing the PLL loop bandwidth and increasing the pre-correlation integration time reduces the susceptibility of the PLL to interference that causes a decrease in the carrier-to-noise ratio.

Section A.1.

Narrowing the noise bandwidth of the PLL was successfully implemented and studied by *Manz et al.* (2000) for implementation in WAAS Reference Stations. The authors decreased the noise bandwidth for a third-order PLL from 15 dB-Hz to 3 dB-Hz and obtained improvements of over 12 dB on GPS satellite signals and a 5 dB improvement on WAAS signals. The difference in improvement between GPS and WAAS is that the pre-correlation integration time is only 2 ms for WAAS, as opposed to 20 ms for GPS. The WAAS Reference Stations have the added benefit of being stationary, which helps to reduce jerk-induced tracking error (θ_e), as well as reducing vibration-induced oscillator phase noise (σ_v^2, explored in the Appendix, Section A.2). The WAAS Reference Stations also implement cesium atomic frequency standards, which are substantially more stable than any crystal oscillator over long time intervals. The cesium standard significantly decreases the

contribution of Allan Deviation oscillator phase noise (σ_A), rendering the receiver substantially less susceptible to loss of lock at low carrier-to-noise ratios. This concept will be further explored below.

For some applications, it is not desirable to reduce the noise loop bandwidth to such low levels because the receiver will likely lose carrier tracking in high dynamic situations. In this case, an adaptive PLL could be implemented where the noise loop bandwidth is only decreased for low carrier-to-noise ratios. This idea was explored by *Humphreys et al.* (2005) as applied to ionospheric scintillation and is readily applied to solar radio bursts. By reducing the noise loop bandwidth only during periods of decreased carrier-to-noise ratio, the receiver maintains lock on the satellite, then increases the noise loop bandwidth once the interference has gone away. The deepest fades caused by solar radio bursts may only last for a few tens of seconds. The interval may be short enough that the receiver maintains lock, or at least will quickly re-obtain lock once it is lost. Otherwise, during strong periods of interference, the receiver may indicate the interference at which point high-dynamic maneuvers should be avoided.

Receiver Clock Noise and Allan Variance

In most GPS receivers, cheap, temperature-compensated crystal oscillators (TCXO) are employed. In higher grade survey and scientific receivers, ovenized crystal oscillators (OCXO) are utilized for greater clock stability and improved carrier phase measurements. The ultimate clock available to a GPS receiver is an atomic frequency standard that provides the cleanest phase measurements over long intervals.

The phase error introduced by the clock is related via the Allan variance, $\sigma_A(\tau)$,

which is determined for the oscillator over a gate time, τ. This parameter is given by oscillator manufacturers for gate times of 1 second, or it can be readily calculated from phase noise measurements (*Allan*, 1987). The Allan variance (a unit-less quantity) is a measurement of stability in oscillators, and is determined empirically as one half of the time average (τ) of the squares of the differences between successive readings (δy) of the frequency deviation sampled over the sampling period (*Allan*, 1987):

$$\sigma_A^2(\tau) = \frac{1}{2}\langle(\delta y)^2\rangle \tag{6.4}$$

The Allan variance-induced oscillator jitter (1-σ) due to the receiver clock for a third-order PLL is calculated from the Allan variance as (*Ward*, 1996a):

$$\theta_A = 160\frac{\sigma_A(\tau)f_L}{B_L} \text{ degrees} \tag{6.5}$$

where f_L is the operating frequency (L1 or L2), and B_L is the noise-loop bandwidth of the PLL. Note that the Allan variance is evaluated at the noise bandwidth of the PLL.

Figure 6.2 illustrates how the phase noise contributes to the error in measuring carrier phase. Each line is plotted for a constant Allan variance as a function of loop noise bandwidth. For a given oscillator, the Allan variance differs for varying loop noise bandwidths. The goal is to decrease the loop phase noise as much as possible to eliminate interference effects, as shown in the previous section. When considering this technique, however, the phase error due to clock phase noise is inversely proportional to the noise loop bandwidth. Care must be taken to choose an appropriate clock to minimize the phase error at narrow bandwidths.

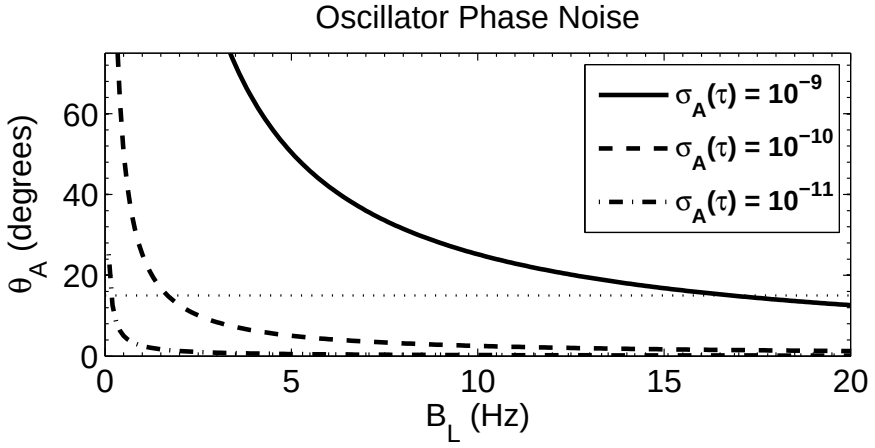

Figure 6.2: The phase error due to clock phase noise as a function of PLL loop bandwidth for a third-order PLL operating on the L1 frequency. Each line represents a line of constant Allan variance, which is a measure of phase noise in an oscillator. The horizontal dotted line represents the maximum tolerable error of 15°. Notice that a more stable clock results in decreased phase error, however, all clocks have a dramatic increase in phase error at narrow bandwidths.

Typically, GPS receivers utilize TCXO's as their frequency standard because they are cheap, small, low power, and very stable over short time durations. In fact, TCXO's and OCXO's are more stable than cesium atomic frequency standards over short time durations. TCXO's, however, exhibit large drifts over longer time intervals.

As an example, a Rakon TXO4080 oscillator[3] is the TCXO that is found on the Cornell ScintMon receivers and has a stability at a bandwidth $B_L = 1$ Hz of at best $\sigma_A(1) = 2 \cdot 10^{-10}$, corresponding to phase errors of $\theta_A = 5°$ for a third-order PLL at L1. In comparison, a SpectraTime LCR-900 Rubidium atomic frequency standard[4] has a $B_L = 1$ Hz (a long interval of $\tau = 1$ s) stability of $\sigma_A(1) = 1 \cdot 10^{-13}$, corresponding to a phase error of $\theta_A = 0.002°$. Atomic frequency standards have a much clearer advantage over long-term intervals, and quartz-oscillators steered by atomic standards are commonly used. Naturally, the cost of a Rubidium standard

[3]http://www.rakon.com/models/browse-model?model_id=89&model_type=O
[4]http://www.spectratime.com/product_downloads/lcr_spec.pdf

(several thousand dollars) is significantly more expensive than a TCXO or OCXO (less than a few hundred dollars per unit in most cases for OCXO's).

Again, the stability of the WAAS Reference Stations during the solar radio burst is in part due to the fact that they employ Cesium atomic frequency standards (*Townsend et al.*, 1995). The clock stability, coupled with the very narrow noise loop bandwidth of the third-order PLL, and the lack of receiver-satellite dynamics aids the receiver through the interference from the solar radio burst.

Frequency Lock Loop

Phase Lock Loops operate by accurately tracking the phase of the incoming carrier signal. They are only necessary when precise carrier phase or velocity measurements are needed, but they are particularly susceptible to dynamic stress. Contrary to PLL's which generate an estimate of the phase of the incoming signal, a Frequency Lock Loop (FLL) performs carrier wipeoff by replicating the approximate frequency (*Ward*, 1996a). FLL's are less susceptible to dynamic stress and are frequently used in the acquisition phase of signal tracking because they can acquire a signal more easily than a PLL. Furthermore, they are negligibly susceptible to clock phase noise and vibration-induced oscillator jitter. Thus, the FLL is only sensitive to thermal noise frequency jitter and dynamic stress error.

According to *Ward* (1996b), the 3-σ tracking error for the FLL cannot exceed 90° in a pre-correlation integration time interval, T. The 1-σ tracking error must be less than 30°, which is equivalent to 1/12 of a cycle. The 1-σ FLL tracking threshold is given as:

$$\sigma_{FLL} = \sigma_{\phi_{FLL}} + f_e \leq \frac{1}{12T} \text{ Hz} \tag{6.6}$$

where $\sigma_{\phi_{FLL}}$ is the track error due to thermal noise, and f_e is the frequency error caused by dynamic stress.

The FLL tracking loop error due to thermal noise is approximately:

$$\sigma_{\phi_{FLL}} = \frac{1}{2\pi T}\sqrt{\frac{4B_L}{C/N_0}S_L} \text{ Hz} \tag{6.7}$$

where $T = 1/B_P$ is the pre-correlation integration time interval, B_L is the loop noise bandwidth, C/N_0 is the carrier-to-noise ratio in Hz, and S_L is the squaring loss, which is given as:

$$S_L = 1 + \frac{1}{TC/N_0} \tag{6.8}$$

Equation 6.7 is independent of FLL loop order and is applicable to any PRN sequence. Figure 6.3 illustrates the thermal noise performance of the FLL. The lines shown in the plot are for the same loop parameters as the first three lines plotted in Figure 6.1 (the PLL). Notice that for a given noise bandwidth, B_n, and pre-correlation integration time, T, the PLL and FLL have comparable performance. Once again, narrowing the noise bandwidth improves immunity to low carrier-to-noise ratios, and increasing the pre-correlation integration time also improves immunity.

The advantage of the FLL over the PLL is that the FLL has significantly better dynamic stress performance than the PLL (please refer to the Appendix, Section

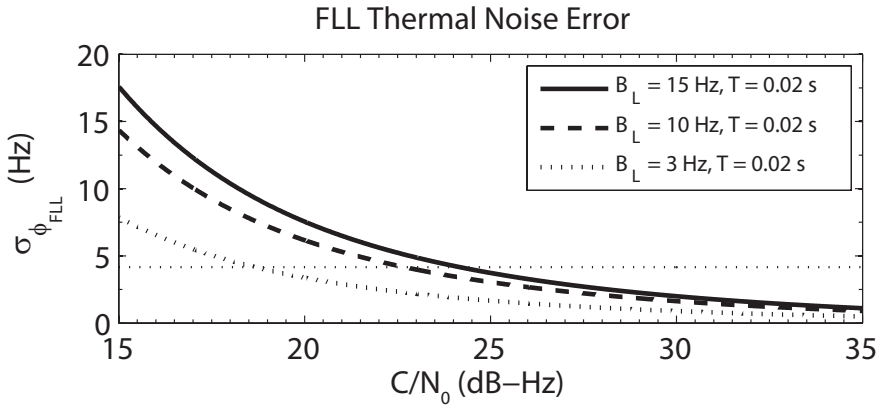

Figure 6.3: The tracking error in Hz due to thermal noise as a function of carrier-to-noise ratio. The lines in this plot are the same as the first three lines shown in Figure 6.1. The horizontal dashed line at 4.16 Hz represents the tracking threshold for an FLL operating with $T = 0.02$ s.

A.1). The increased dynamic stress performance also corresponds to an advantage during acquisition of a satellite signal where the FLL outperforms the PLL (*Ward*, 1996a). The FLL, therefore, can operate at significantly lower bandwidths than the PLL, allowing for more tolerance of noise. At the very least, a well designed receiver should implement an FLL as a fall-back when the PLL loses carrier-track, either as a result of low carrier-to-noise ratio or for re-acquisition. A good receiver will also implement an FLL when performing initial acquisition with a hand-off to a PLL once the code offset, Doppler shift, and carrier phase of the satellite have been ascertained.

Before leaving the subject of tracking loop modification, it is worth noting that there exists a type of tracking loop known as the 'FLL-Assisted-PLL' carrier tracking loop which marries the advantages of the FLL (decreased sensitivity to dynamic stress) and the PLL (accurate carrier phase measurements) and provides modest increases in tracking threshold over the standard PLL. The interested reader is referred to *Ward* (1998).

6.2.2 Bit Prediction

Squaring loss (Equation 6.3) is problematic because it does not allow weak signals ($C/N_0 < 25$ dB-Hz) to be tracked. Squaring loss arises because the incoming GPS signal must be squared to eliminate the 180° phase transitions associated with data-bit transitions. The squaring of the signal also squares the noise, resulting in decreased performance of the tracking loop.

Another problem that arises with data-bit transitions is that the length of time for integrations is limited by the data-bit rate. As was shown in Figures 4.4 and 6.1, increasing the integration time, T, allows for even smaller carrier-to-noise ratios to be tracked.

The current navigation message transmitted on the L1 C/A and L1 and L2 P(Y) code is a highly redundant 50 BPS signal (*IS-GPS-200D*, 2004). The satellite ephemeris are only updated every two hours. Three of every five subframes transmitted in the data message contain the satellite ephemeris, and few bits change from one subframe to the next (e.g. the Z-count). The other two subframes contain specialized messages including the satellite almanac that cycle through 25 pages every 12.5 minutes and are updated every 3.5 days *IS-GPS-200D* (2004). The result is a highly repetitive signal, for which the bitstream is easy to predict.

Predicting the bitstream and applying the proper bit phase allows for the receiver to base-band mix the signal so that the signal does not have to be squared within the PLL. As a result, squaring loss has been eliminated from Equation 6.2, thus significantly reducing the tracking error in Equation 6.1 allowing for smaller carrier-to-noise ratios to be tracked.

With the removal of data-bit transitions, the receiver is also capable of inte-

grating over intervals that are longer than the data-bit period (currently 20 ms for the navigation message). Longer integrations accumulate the weak signal while the noise does not, thus facilitating detection of the weak signal. Long integration times, however, require a very stable clock so as not to introduce excessive clock noise into the tracking variance (Equation 6.1), ultimately resulting in a loss of carrier tracking. Another limiting factor in long integration time is that the receiver-satellite dynamics must be within tracking limits over the integration time interval. Integration of the GPS receiver with an IMU would alleviate dynamic stress errors over long integration periods.

Collectively these schemes to increase the GPS receiver sensitivity are two of many techniques used in a process known as Assisted GPS (AGPS). AGPS has been successfully implemented for Enhanced-911 (E911) services on cell phones, which provide emergency responders with the location of the caller. According to *Weill* (2006), these techniques can improve the sensitivity of GPS receivers down to -155 to -165 dBm. Recall from Table 3.3 that the received power at moderate elevations (antenna gain of 0 dB) is around -156.1 dBW = -126.1 dBm. These improvements increase the sensitivity of the receiver down to carrier-to-noise ratios in the neighborhood of 10 - 20 dB-Hz!

Typically the AGPS correction parameters are transmitted from a reference station to a mobile receiver such as a cell phone. This method is employed because a cell phone may be located in weak signal areas, such as inside a building, and also has limited resources (e.g. processing power) to determine a navigation fix. Bit prediction could easily be accomplished by a commercial GPS receiver and would require software and possibly hardware modifications. Future signals are planned for GPS that will include data-less pilot signals, exactly for the purpose

of signal acquisition and tracking in weak-signal environments. The data-less pilot signals also allow the 1-σ tracking error threshold to be increased to 30° because cycle-slips are now considered over 360° as opposed to 180°. The future signals are discussed in Section 6.2.6.

6.2.3 External Aiding Enhancements

External aiding enhancements are measurements from sensors other than GPS that help the receiver to navigate. These external measurements are particularly useful because they are usually not affected by an interfering source. According to *Ward* (1996b), examples of navigation sensors that have been integrated with GPS are inertial measurement units (IMUs), Doppler radar, and air speed, baro altimeter, and magnetic compass sensors. The aiding that these instruments provide varies on their accuracy, but at the very least they can provide velocity, altitude, and heading for periods when the GPS receivers are unable to navigate. If the measurements are more tightly-coupled with the navigation filter and tracking loops, they can actually help the receiver maintain lock on the GPS signals.

IMUs are most commonly found in GPS receivers because they are immune to interference and are highly accurate over short time-spans. Their main disadvantage is that they must be initialized and are prone to drift, but the accuracy of the GPS receiver's navigation solution can be coupled with the IMU to initialize and calibrate the device. In turn, the stability of the IMU over short periods of time can aid the carrier tracking loops in a tightly coupled system. The IMU provides accurate, independent velocity measurements that can be included in the navigation filter to provide accurate velocity aiding. This effectively removes dynamics (e.g. acceleration and jerk stress) from the receiver tracking loops. As a result, the IMU-

aided tracking loops are operated at extremely narrow noise bandwidths, which as shown in Section 6.2.1, greatly increases the resiliency of the GPS receiver to weak carrier-to-noise ratio environments. Moreover, if the interfering signal eventually over-powers the GPS receiver, then the IMU can continue to navigate through the interference. Once the interference has passed, the measurements from the IMU are used to rapidly re-acquire the GPS signals.

Inertial measurement units are not cheap, and their integration into a GPS receiver typically requires complex Kalman filter solutions. Furthermore, the receiver is usually built with IMU capability because low-level access to the tracking loops is required. As a result, tightly-coupled GPS receivers and IMUs are restricted to use in military and high-end commercial equipment.

6.2.4 Pre-correlation Filters and Digitization

The pre-correlation filter is the band-pass filter placed immediately prior to digitization that band-limits the signal to avoid aliasing. The pre-correlation filter also eliminates unwanted out-of band signals. For an L1 C/A code receiver, the single-sided bandwidth of the pre-correlation filters is typically around 1 MHz. As a result, only the main-lobe of the C/A code spectrum (with bandwidth of 2.046 MHz) is digitized. The higher-order square-wave harmonics of the C/A code sequence are cut-off, resulting in correlation losses. According to *Van Dierendonck* (1996), in the presence of wideband noise, the L1 C/A code receiver will exhibit approximately 0.5 dB of correlation loss. If the single-sided bandwidth is raised to 4 MHz, the loss decreases to 0.2 dB.

The disadvantage of having a wider pre-correlation filter is that the receiver is

introducing a larger bandwidth within which an interfering signal could fall. The trade-off between the potential advantage of processing gain and decreased implementation loss must be weighed against the possibility of the negative effects of the inclusion of potential in-band interfering signals when using wider pre-correlation filters. Additionally, wider pre-correlation filters necessitates faster sampling of the analog signal to avoid violating the Nyquist-Shannon sampling theorem. A higher sampling rate is costly because it requires faster processing speeds, which primarily increases power consumption.

The WAAS receivers utilize an 18 MHz front-end, which is substantially wider than typical L1 C/A code only receivers (private communication with Patrick Reddan). The WAAS receiver is a semi-codeless dual-frequency receiver and so the wide bandwidth is needed to accurately track the P(Y) code. The wider bandwidths appear to have helped the WAAS receiver to adequately navigate through the solar radio burst, but only if the bandwidth of the radio bursts were relatively narrow as compared to the front-end bandwidth.

The number of bits used in the digitization of the analog GPS signal also causes implementation losses within the receiver and is tightly coupled to the pre-correlation filter bandwidth. There are many considerations that must be weighed when digitizing the signal, including whether to include a zero-level crossing. The GP2015 chip-set by Zarlink, for example, does not contain a zero-level crossing, which is detrimental when the incoming signal lies between thresholds. This effect is offset in the GP2015 by the inclusion of the automatic gain control (AGC), which will keep the signal at the appropriate levels to illuminate the analog to digital converter. According to *Van Dierendonck* (1996), the best-case implementation loss for 1-bit quantization (hard-limiting) is 3 dB, 0.73 dB for 2-bit quantization,

approximately 0.21 dB for 3-bit quantization, 0.068 dB for 4-bit quantization, and about 0.032 dB for 5-bit quantization. The cases are presented for pre-correlation filters of single-sided bandwidth equal to the chipping rate of the PRN sequence. When a wider pre-correlation filter is used, the values for all schemes drop slightly.

Again, increasing the number of bits in the receiver requires more power, and is especially costly in hardware complexity. The net results to be expected from using, for example, a 5-bit digitization scheme with a pre-correlation bandwidth five times wider than the chipping rate results in a minimal loss of approximately 0.11 dB, as opposed to a 5-bit system utilizing a pre-correlation filter equal to the chipping rate, which suffers about 0.5 dB of loss.

6.2.5 Nulling Antenna Arrays

Nulling antenna arrays introduce a steerable null into the cumulative antenna gain pattern. When the null is steered in the direction of the interfering source, the intensity of the source can be significantly reduced. Nulls are created by the inclusion of two or more antennas whose signals are added together with an appropriate complex weight to steer the null in the direction of the interference.

The following simplified analysis is from *Misra and Enge* (2006), which is developed from *Bauregger* (2003) and *Rounds* (2004). Suppose two antennas are separated by a distance $d = \lambda/2$, shown in Figure 6.4[5]. The continuous-time representation of the down-converted signal within the receiver from the right-hand antenna is represented by:

[5]Choosing other values of $d < \pi/2$ results in wider nulls, and values of $d > \pi/2$ results in parasitic nulls.

$$S_R(t) = \sqrt{P_C}C(t)D(t)\mathrm{e}^{j(2\pi\Delta f_D t+\Delta\phi)} + \sqrt{P_J}\mathrm{e}^{j(2\pi f_J t+\Delta\phi_J)} \tag{6.9}$$

where the left-hand term is the GPS signal as defined in Equation 3.1. Δf_D is the difference between the true Doppler shift f_D, and the receiver generated Doppler shift $\hat{f}_D$, and $\Delta\theta$ is the difference between the received phase and the receiver-generated phase. The term on the right is the interference term, which for the sake of this argument is represented simply by a cosine wave. Δf_J and $\Delta\theta_J$ are the corresponding Doppler and phase offsets for the interfering signal.

From the left-hand antenna, the arriving signals are similar, except that there is a phase offset of the satellite and interfering source caused by the additional signal path length, $d\cos(\theta)$ for each satellite, and $d\cos(\theta_J)$ for the interfering source. The base-band signal within the receiver from the left-hand antenna is:

$$S_R(t) = \sqrt{P_C}C(t)D(t)\mathrm{e}^{jkd\cos(\theta)}\mathrm{e}^{j(2\pi\Delta f_D t+\Delta\phi)} + \sqrt{P_J}\mathrm{e}^{jkd\cos(\theta_J)}\mathrm{e}^{j(2\pi f_J t+\Delta\phi_J)} \tag{6.10}$$

where $k = 2\pi/\lambda$ is the wave number.

Multiplying Equation 6.10 by a complex weight, $W_L = -e^{(-jkd\cos(\theta_J))}$ and adding it to Equation 6.9 allows the interfering signal (and satellite signals lying within the null pattern) to be eliminated. Recall that the signal collected by the antenna is dominated by noise. An interfering source raises the total power of the noise-floor, so a search algorithm to minimize the weighted sum $S_R(t) + W_L S_L(t)$ is easily implemented. The resulting signal is thus:

$$S_R(t) + W_L S_L(t) = \sqrt{P_C}C(t)D(t)\mathrm{e}^{j(2\pi\Delta f_D t+\Delta\phi)} \times (1 - \mathrm{e}^{jkd(\cos(\theta)-\cos(\phi_J))}) \tag{6.11}$$

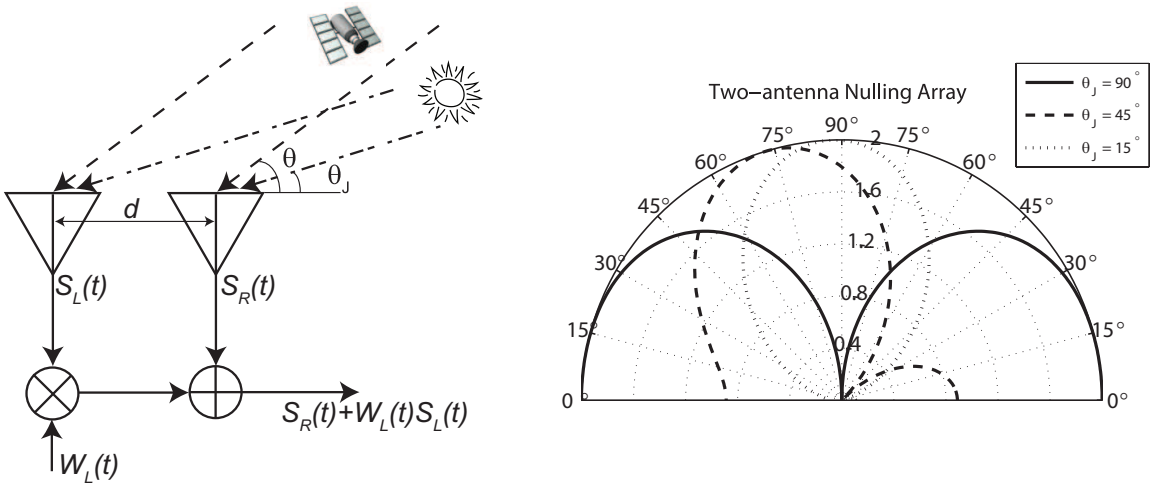

Figure 6.4: The left panel illustrates the two-antenna nulling setup. The right panel illustrates the null created for various interference elevation angles for $d = \lambda/2$.

and the amplitude of the received signals is the term:

$$1 - \mathrm{e}^{jkd(\cos(\theta)-\cos(\theta_J))} \tag{6.12}$$

and is shown in the right panel of Figure 6.4.

In the case of solar radio bursts, the sun is the interfering source and its elevation and azimuth as a function of time is calculable (e.g. *Reda and Andreas* (2005)). During a period of solar activity, the null could easily be steered to track to the sun. Unfortunately, any satellite signals that fall within the null will also be attenuated. GPS antennas typically receive GPS signals of moderate to high-elevation with at least 45 dB-Hz C/N_0. Assuming an interference source at $\theta_J = 45°$ and an antenna spacing of $d = \lambda/2$, a satellite laying within approximately 2.5° of the null bore-sight will suffer greater than 10 dB of attenuation (i.e. the satellite signal will be received with a carrier-to-noise ratio of at most 35 dB-Hz). For GPS, the likelihood of the null containing a satellite signal is quite small and is a small price

to pay in the presence of intense solar radio bursts.

Adaptive antenna arrays typically include 7 to 9 antennas, allowing for sharper and multiple nulls (*Rounds*, 2004). In general, an array of N antennas can have N-1 steerable nulls. Although the position of the source (the sun) is not changing azimuth or elevation rapidly, the intensity and primary frequency of the solar radio burst may be changing rather quickly. The weight, W_L, added to the signals should contain a gain factor to properly eliminate the interfering signal. For interference sources laying within ±20 MHz of GPS L1 (the bandwidth of the GPS signals), the null pattern remains essentially unaltered if the two-antenna nulling array spacing is fixed at $d = \lambda/2$ (about 9.515 cm for GPS L1).

Adaptive antenna arrays are expensive both in cost and implementation. The receivers that are capable of processing these signals must have RF front-ends for each incoming antenna, and specialized hardware to add the signals appropriately. Software receiver approaches would still require multiple front-ends, but the addition of the signals could take place in a Digital Signal Processor, reducing hardware complexity at a cost of increased signal processing. Nulling antenna arrays, however, can offer 25 dB or more immunity from broadband interference sources (*Ward*, 1996b), arguably making them the best defense against solar radio bursts.

6.2.6 Satellite and Signal Augmentation

Perhaps the ideal way to overcome solar radio bursts is to increase the satellite transmission power. In fact, the minimum received power that is cited in *IS-GPS-200D* (2004) is the specification for the end of the satellite's life. In most cases,

the received signal power is significantly higher when the satellite is first launched. Increasing the satellite power requires new satellites to be designed and launched, which is extremely expensive in cost, design time, and replenishment time.

Already, the transmitted signal power from the GPS satellites is improving. The current generation of Block II/IIA/IIR satellites transmit with a minimum signal strength of -161.5 dBW on L1 P(Y), -158.5 dBW on L1 C/A, -164.5 dBW on L2 P(Y). The generation of satellites currently being launched, Block IIR-M/IIF, have increased the minimum guaranteed signal powers of L2 P(Y) to -161.5 dBW and have added the L2C (Civilian) signal at -160.0 dBW and M-code (Military) signal at -158.0 dBW. The increased power of the L2 P(Y) signal renders the military less susceptible to interference effects. The L2C signal provides a civilian access signal at L2 which will allow for ionospheric corrections as well as greater frequency diversity.

Frequency diversity reduces the susceptibility of the receiver to interference because the solar radio burst may have significantly different intensity at L1 or L2, as occurred on 13 December 2006 (Figure 5.18) or 14 December 2006 (Figure 5.20). A civilian dual-frequency receiver could recognize that one frequency is less affected than another, and apply the phase or Doppler measurements from the carrier tracking loop to aid the carrier tracking loop of the more affected frequency. When one tracking loop is carrier-aided in this manner, the tracking loop for the weaker signal can have a much narrower PLL (or FLL) noise bandwidth, which as shown in Section 6.2.1, greatly improves the receiver's ability to cope with weak carrier-to-noise ratios. In fact, the WAAS receivers employ such a technique whereby the L2 tracking loops are aided by measurements from the L1 tracking loops. The tracking loop noise bandwidth on L2 is substantially decreased allowing

the signal to be tracked more easily (private communication with Pat Reddan). Moreover, if lock is eventually lost, the signal is rapidly reacquired with the Doppler measurements from the unaffected signal.

The L2C signal is divided into two components. One signal is known as the L2C-M signal, which is a BPSK, medium-length code of 20 ms period at a chipping rate of 511.5 Kilo-chips per second. The L2C-M signal contains a data signal which has Forward Error Correction (FEC) encoded by a rate 1/2 convolutional code. The resulting symbol rate is 50 symbols per second, with an effective bit rate (after deconvolution) of 25 Bps. The other L2 signal is known as the L2C-L. The L2C-L code has a 1.5 s period, also at a chipping rate of 511.5 Kilo-chips per second and is a data-less pilot signal. Every chip of the L2C-M code is time multiplexed with a chip of the L2C-L code, and so the effective bandwidth of the L2C-M and L2C-L signals is the same as the L1 C/A code. The ranging precision of the three signals, therefore, is equal. However, the L2C-M code has a code period equal to the symbol period (20 ms), which allows for simplified weak-signal acquisition since the receiver does not run the risk of crossing bit-boundaries during an integration period. Furthermore, since the L2C-L code is data-less, squaring loss is no longer a concern and long integrations can be performed in weak signal environments, as discussed in Section 6.2.2.

The Block IIR-M, IIF and III GPS satellites are being launched with new military signals, known as the M-code. With a nominal power of -158.0 dBW, the Block III GPS satellites will allow for spot-beam technology, where the transmission power of the M-code can be further increased by 20 dB over limited geographical areas (*Barker et al.*, 2000). The M-code is a Binary Offset Carrier (BOC(10,5)) signal with a PRN sequence chipping at 5.115 MCps with a sub-carrier chipping

rate of 10.23 Mcps. BOC codes are simply PRN sequences modulated by a square-wave subcarrier of higher chipping rate than the PRN sequence (*Betz*, 2002). The result is that the majority of the energy contained in the spectrum is not centered at L1 or L2 but rather at an offset of approximately ± 10 MHz (see Figure 6.5). The spectrum offers the advantage that it can be superposed on top of the existing GPS signals without interference. The BOC codes provide no increased protection to wide-band RFI because the signal is still a data-modulated PRN sequence subject to the same squaring losses as the L1 C/A or L1/L2 P(Y) codes. The wider bandwidth as compared to the current BPSK PRN sequences implies that a much larger bandwidth is required to track these signals, which may make them more prone to solar radio burst interference. BOC codes, however, provide better ranging accuracy at weak carrier-to-noise ratios because the auto-correlation peak is narrower (*Betz*, 2002).

The next generation of GPS satellites, known collectively as GPS-III, will transmit a new signal at 1.117645 GHz known as L5 with a BPSK PRN sequence clocked at 10.23 Mcps primarily for aviation purposes, but publicly accessible. L5 falls in the Aeronautical Radio Navigation Service (ARNS) frequency allotment, which means that the band is protected and solely reserved for the use of aviation. The L5 signal will be received with a minimum power of -157.9 dBW, as opposed to -158.5 dBW for L1 C/A and -160.0 dBW for L2C (*IS-GPS-705*, 2005). The 0.6 dB higher transmission power will slightly help when interference from solar radio bursts is present. The PRN sequence chipping rate (10.23 Mcps) is significantly higher than the codes found on L1 C/A or on the L2C signal (1.023 Mcps), which reduces the number of auto- and cross-correlation side-lobes allowing for easier acquisition of the signal. The longer signal and smaller chip period also results in sharper auto-correlation peaks, increasing ranging performance over the L1 C/A

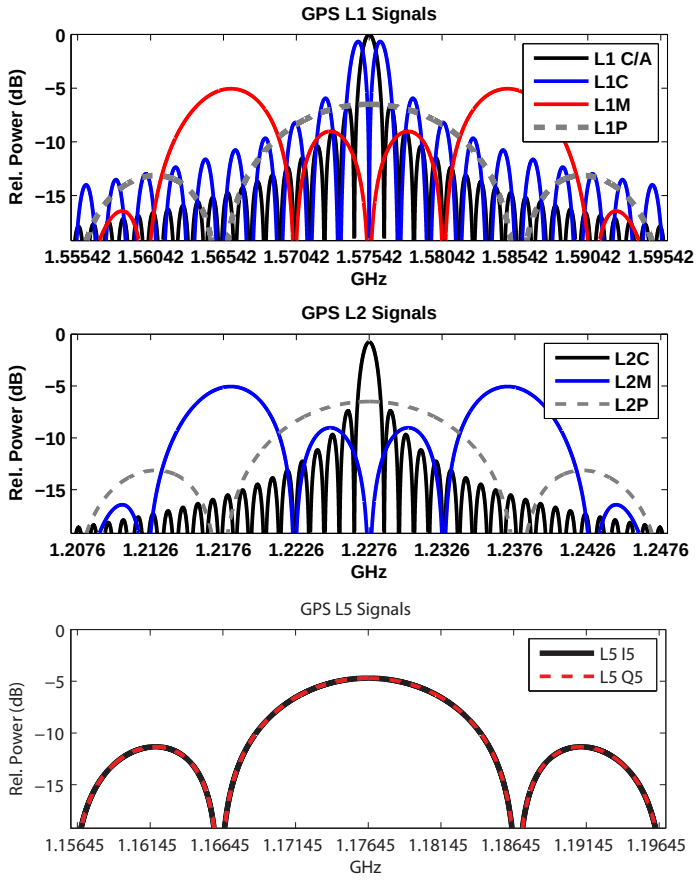

Figure 6.5: From top to bottom, each panel represents the L1, L2, and L5 signals, respectively, as is planned for the Block III GPS satellites. Solid lines represent in-phase signals, while dashed lines represent quadrature signals. In each plot, the power is relative to the L1 C/A code.

code. The larger bandwidth as compared to the L1 C/A code also results in more processing gain, which aids in overcoming narrow-band interference[6], providing modest improvements to all forms of interference. Lastly, the quadrature component of the L5 sequence is data-less, which eliminates squaring loss and allows for longer integration times in weak signal environments.

[6]Processing gain is defined as the ratio $PG = \frac{B_{code}}{B_{despread}}$, where B_{code} is the code chipping rate, and $B_{despread}$ is the equivalent bandwidth after code stripping (usually 1 Hz). The processing gain for the L1 C/A code is 60 dB, where as the L1/L2 P(Y) or L5 processing gain is 70 dB, corresponding to tolerances of roughly 30 dB (40 dB) more power of narrow-band interference than the power contained in the L1 C/A (L1/L2 P(Y)) code.

In addition to the new L5 signal, there are plans to augment the L1 signal with an L1C (civilian) signal which will be a BOC(1,1) signal (*Rushanan*, 2007). The L1C, similar to the L2C, will be time multiplexed and will reside on the L1 in-phase portion of the spectrum. The signal will have a 10 ms period with a 1.023 MCps Weill sequence. Weill sequences have several advantageous properties which will help to minimize inter-signal interference. The $L1C_D$ signal (containing data) will be transmitted at nominally -163.0 dBW, whereas the data-less pilot signal, $L1C_P$, will be transmitted at -158.25 dBW (*IS-GPS-800*, 2007). The L1C BOC signal chipping rate, code-period, and pilot signal will help to increase pseudorange precision. All the signal augmentations discussed herein are summarized in Table 6.1. The amplitude spectrums of the various signals transmitted by the Block III satellites are shown in Figure 6.5.

Another improvement that can be made is to increase the number of transmitting satellites. Currently, the constellation is guaranteed at nominally 24 satellites with 95% availability over any day, with at least 21 satellites transmitting a navigation messages over 98% of a year (*GPS SPS Performance Standard*, 2001). From 2003 - 2006, there was an average of 28 operational satellites with 27 - 29 active satellites on orbit over the same period (*Lavrakas*, 2006). As of Fall 2007, there are 31 operational satellites in orbit. The large number of operational and transmitting satellites has resulted in significant improvements to precision, since the DOP decreases with the addition of more satellites. Since the satellites are more-or-less randomly spaced throughout the sky for users at low to midlatitudes, the addition of satellites naturally results in an increased chance that there will be more medium- to high-elevation satellites. As was shown repeatedly in the text, a receiver is significantly less likely to drop medium- to high-elevation satellites because their received power is higher than low-elevation satellites.

Table 6.1: Summary of GPS civil and military signals. ^{I}-This signal is found on the in-phase component. ^{Q}-This signal is found on the quadrature component. *-The military signals are encrypted and not publicly available. †-The 50 Hz data message with only single bit-error detection. ‡-The data are 1/2 rate FEC, resulting in a data rate half the symbol rate. $^{\flat}$-The Military M-code is a BOC signal, as opposed to the other signals that are BPSK. In the chip rate column, the given rates are the sub-carrier frequency and the spreading frequency, respectively. In literary nomenclature, this results in a BOC(10,5) signal for the M-code and a BOC(1,1) signal for the L1C signal. $^{\sharp}$-On Block III satellites the M-code power can be dynamically increased up to -138.0 dBW. ?-The code-period and navigation message on the M-code is not known.

GPS Block	Center Freq.	Signal	Chip Rate (MCps)	Code Per.	Power (dBW)	Symbol (sps)	Purpose
I	L1	C/A^{I}	1.023	1 ms	-158.5	$50^{\dagger}$	Civ.
IIA	L1	$P(Y)^{Q}$	10.23	1 wk	-161.5	$50^{\dagger}$	Mil.*
IIR	L2	$P(Y)^{Q}$	10.23	1 wk	-164.5	$50^{\dagger}$	Mil.*
IIR-M	L1	C/A^{I}	1.023	1 ms	-158.5	$50^{\dagger}$	Civ.
	L1 & L2	$P(Y)^{Q}$	10.23	1 week	-161.5	$50^{\dagger}$	Mil.*
	L2	C-M^{I}	1.023	20 ms	-163.0	$50^{\ddagger}$	Civ.
	L2	C-L^{I}	1.023	1.5 s	-163.0	Pilot	Civ.
	L1 & L2	$M^{I\flat}$	10.23, 5.115	?	-158.0	?	Mil.*
III	L1	C/A^{I}	1.023	1 ms	-158.5	$50^{\dagger}$	Civ.
	L1	$C_{P}^{I\flat}$	1.023, 1.023	10 ms	-158.25	Pilot	Civ.
	L1	$C_{D}^{I\flat}$	1.023, 1.023	10 ms	-163.0	$100^{\ddagger}$	Civ.
	L1 & L2	$P(Y)^{Q}$	10.23	1 week	-161.5	$50^{\dagger}$	Mil.*
	L2	C-M^{I}	1.023	20 ms	-157.0	$50^{\ddagger}$	Civ.
	L2	C-L^{I}	1.023	1.5 s	-157.0	Pilot	Civ.
	L1 & L2	$M^{I\flat}$	10.23, 5.115	?	$-158.0^{\sharp}$	?	Mil.*
	L5	$I5^{I}$	10.23	1 ms	-157.9	$100^{\ddagger}$	Civ.
	L5	$Q5^{Q}$	10.23	1 ms	-157.9	Pilot	Civ.

Furthermore, the resurgence of Russia's GLONASS and the addition of other GNSS systems, such as the European Union's Galileo and the Chinese' Beidou (Compass), will result in hybrid receivers that take advantage of and will assimilate all the available navigation signals.

For example, with a constellation of minimally 24 satellites for GPS and min-

imally 30 satellites for Galileo (which are in similar orbits), the typical number of visible satellites will easily double. Of course, the cost and complexity of receivers will increase, but this is a small price to pay for more accurate, guaranteed service. Hybrid GPS/GLONASS receivers are currently available on the market (e.g. NovaTel OEMV2[7]), and even hybrid GPS/Galileo receivers are available (e.g. Septentrio AsteRx1[8]).

6.3 Estimating Solar Radio Burst Power from GPS Measurements

The FAA WAAS Reference Station network is a chain of highly reliable and robustly maintained identical equipment. The robustness of the equipment yields good data from which estimates of the solar radio burst power can be made. The solar radio burst should illuminate the surface of the Earth equally because, as viewed from the Sun, the Earth occupies about 17.5 arcseconds of angular width. The angle subtended by the Earth is so small that it is unlikely that the solar radio burst would produce a differential radiation pattern on the surface of the Earth (private communication with Dale Gary).

From each FAA WAAS Reference Station, a measure of the carrier-to-noise ratio fade ($\Delta C/N_0$) can be made and inverted to obtain solar radio burst power. Precise values for the gain of the WAAS antenna and the noise figure of the preamplifier are known. The solar radio burst power as observed at each station can be determined from $\Delta C/N_0$, the carrier-to-noise ratio fades.

[7]http://www.novatel.com/Documents/Papers/OEMV2.pdf
[8]http://www.septentrio.com/products_asterx1.htm

At each Reference Station, three receivers operate in tandem from the same oscillator and receiving antenna. The first step in determining the uniformity of the radiation from the solar radio burst is to determine the consistency of data between the receivers at each reference station.

First, an analysis is conducted to determine if the receiver's automatic gain control operated within the linear mode of operation and if the receiver tracking loops were not unduly stressed. The deepest average fade observed across the three receivers within any of the WAAS Reference Stations at the L1 frequency was about 33.9 dB-Hz (observed at Miami). As shown earlier in this chapter, most receivers can adequately track the L1 C/A code signal down to levels of about 25 dB-Hz before cycle slips and loss of carrier lock occurs. The tracking loops were not unduly stressed. Furthermore, as shown in Figure 6.6, there is no clear correlation between the magnitude of the fade ($\Delta C/N_0$), and the pre-fade value (C/N_0), indicating that the automatic gain control did not saturate. Conversely, the fade on L2 was so deep, in some instances dropping below 25 dB-Hz, that the tracking loops were severely stressed. Only the fades at L1 will be considered for the analysis.

The next step is to determine the consistency of carrier-to-noise measurements across the WAAS receivers within a Reference Station. Unfortunately, each of the three receivers within the Reference Station operates with different tracking parameters to minimize multipath effects and provide continuous measurements under a multitude of conditions. The top panel of Figure 6.7 illustrates the L1 carrier-to-noise ratios as recorded on each of the receivers within the Reference Station at Houston, Texas from PRN 17. Notice that the multipath variations in received carrier power before and after the solar radio burst vary for each receiver,

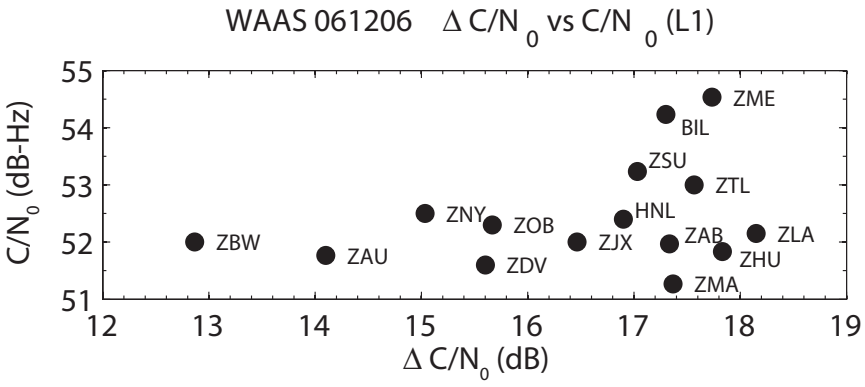

Figure 6.6: The fade ($\Delta C/N_0$) is plotted as a function of the unaffected carrier-to-noise ratio (C/N_0) obtained from a sidereal day earlier for the L1 frequency. There is not a clear correlation between $\Delta C/N_0$ and C/N_0, indicating that the automatic gain control was not saturated.

particularly from 1600 UT to 1700 UT, denoted by the arrow. The large differences in carrier-to-noise ratio are primarily caused by the tracking loop response of each receiver to the local multipath environment.

To determine the solar radio burst flux, however, the quantity of interest is the C/N_0 fade, $\Delta C/N_0$. The bottom panel of Figure 6.7 illustrates the peak fade observed at 1936 UT for each of the three receivers at Houston, Texas at the L1 frequency on PRN 17. Note that there are discrepancies in the measured fade across the three collocated receivers, particularly from receiver 3 and receivers 1 and 2. At this point, multipath has been removed, so the differences are primarily due to noise and the tracking loop response of each receiver to the rapidly changing signal.

To determine the similarity of the observed fades ($\Delta C/N_0$), the variance of $\Delta C/N_0$ must be ascertained for each receiver. Typically the noise variance of the carrier-to-noise ratio measurements is theoretically derived from tracking loop parameters, which in this case are unknown. Therefore, the variance is an empirically derived quantity from $\Delta C/N_0$. The carrier-to-noise ratio fade, $\Delta C/N_0$, is

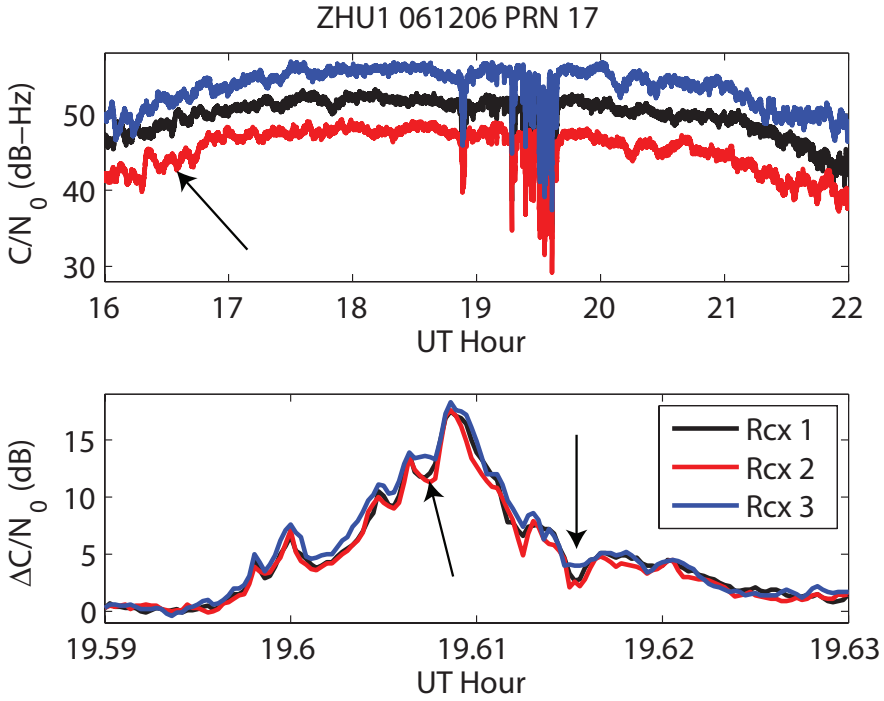

Figure 6.7: Each WAAS Reference Station is actually three receivers operating in tandem. Each of these receivers has slightly different tracking parameters, as is illustrated by the slight differences in the recorded carrier-to-noise measurements due to the tracking loop response to multipath, denoted by the arrow (top panel). The red curve is shifted down by 3 dB-Hz and the blue curve is shifted up by 5 dB-Hz for clarity. The bottom panel illustrates the peak fade measured at each of the three collocated receivers. Note the large discrepancies in $\Delta C/N_0$, particularly between receiver 3 and receivers 1 and 2, denoted by the arrows. All data are from 6 December 2006 at the L1 frequency from the FAA WAAS Reference Station in Houston, Texas.

the difference in the carrier-to-noise ratio measured during the solar radio burst, $C/N_{0_2}(t)$, subtracted from the carrier-to-noise ratio measured a sidereal day before the solar radio burst, $C/N_{0_1}(\tilde{t})$ ($\tilde{t} = t - 86,164.0905$, where 86,164.0905 is the number of seconds in a sidereal day) for the same satellite:

$$\Delta C/N_0(t) = C/N_{0_1}(\tilde{t}) - C/N_{0_2}(t) \text{ dB} \tag{6.13}$$

This process removes the expected variations in received carrier power primarily

due to satellite and receiver antenna gains and multipath.

$C/N_{0_1}(\tilde{t})$ and $C/N_{0_2}(t)$ will have independent noise variances of $\sigma_1^2(\tilde{t})$ and $\sigma_2^2(t)$, respectively. The overall noise variance $\sigma^2_{\Delta C/N_0}(t)$ of Equation 6.13 is:

$$\sigma^2_{\Delta C/N_0}(t) = \sigma_1^2(\tilde{t}) + \sigma_2^2(t) \text{ dB}^2 \tag{6.14}$$

Where $\sigma^2_{\Delta C/N_0}(t)$ is determined empirically over 1,000 one-second samples. An average of the carrier-to-noise ratio for the quiet previous day is also calculated for later use and is denoted $\langle C/N_0(t)\rangle$. The noise variance is calculated empirically from Equation 6.14 because it removes the multipath and expected variations from C/N_0, which would cause the estimates of the noise variance to be skewed if the variance were calculated directly from C/N_0.

With the elimination of multipath, the noise variances of each signal are primarily caused by thermal noise and are gaussian independent variables. The noise variances are also inversely proportional to C/N_0 (see top panels of Figure 6.8), and barring any large differences in the receiver from day to day, will have the same value. During a quiet period when $C/N_{0_1}(\tilde{t}) \cong C/N_{0_2}(t)$ (i.e. with no solar radio bursts), the noise variance of $\sigma_1^2(\tilde{t})$ and $\sigma_2^2(t)$ will be equal because the carrier-to-noise ratios are approximately equal and is calculated as:

$$\sigma_1^2(\tilde{t}) = \sigma_2^2(t) = \frac{1}{2}\sigma^2_{\Delta C/N_0}(t) \text{ dB}^2 \tag{6.15}$$

During the solar radio burst (at t_s), the noise variance of $\sigma^2_{\Delta C/N_0}(t_s)$ is calculated by determining the noise variance $\sigma_1^2(\tilde{t_s})$ for the unaffected, previous sidereal day carrier-to-noise ratio, $C/N_{0_1}(\tilde{t_s})$, and from the noise variance $\sigma_2^2(t_s)$, for the affected

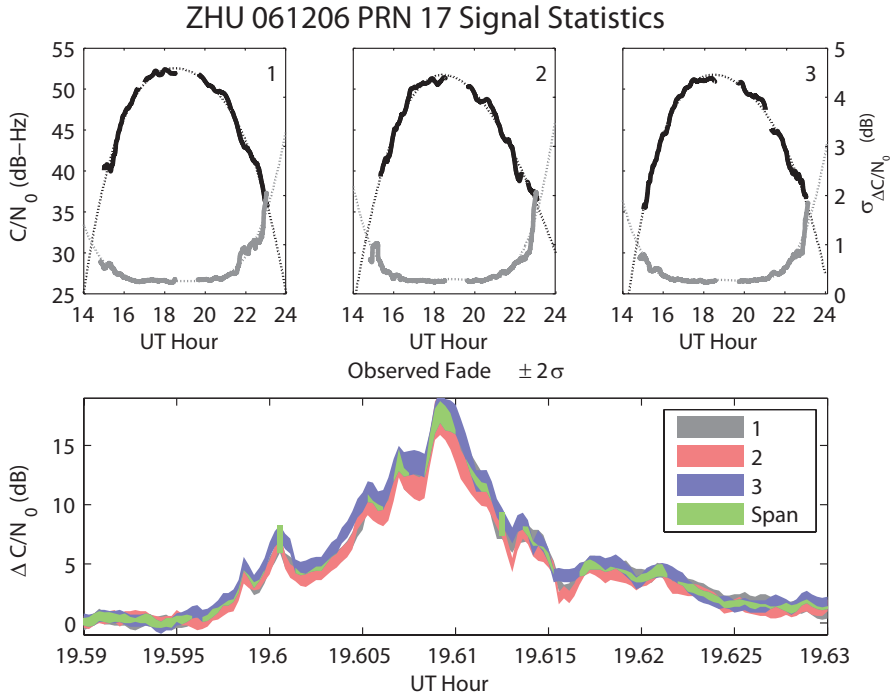

Figure 6.8: The top panels illustrate the mean C/N_0 (solid black line) and variance of $\Delta C/N_0$ (solid gray line) recorded at each of three receivers within the Houston (ZHU) WAAS Reference Station for PRN 17. Clearly the variance of $\Delta C/N_0$ is inversely related to the mean C/N_0. The fourth-order polynomial fits to each line are represented as the dashed line. The bottom panel illustrates fade calculated at each receiver within the Reference Station assuming $\pm 2 - \sigma$ deviation in the expected fade. The green line represents the span of the fade for the three receivers.

carrier-to-noise ratio during the burst, $C/N_{0_2}(t_s)$. First, the noise variance for $\sigma_1^2(\tilde{t_s})$ is found by obtaining $\sigma^2_{\Delta C/N_0}(t)$ when $C/N_{0_1}(\tilde{t_s})$ is equal to $\langle C/N_0(t)\rangle$ during quiet periods. Similarly, $\sigma_2^2(\tilde{t_s})$ is found by obtaining $\sigma^2_{\Delta C/N_0}(t)$ when $C/N_{0_2}(\tilde{t_s})$ is equal to $\langle C/N_0(t)\rangle$ during quiet periods. The variances $\sigma_1^2(\tilde{t_s})$ and $\sigma_2^2(\tilde{t_s})$ are then summed together as in Equation 6.14 to obtain an estimate of $\sigma^2_{\Delta C/N_0}(t)$.

Typically data are available at all values of C/N_0. In the case that a particular value of $\langle C/N_0(t)\rangle$ or $\sigma^2_{\Delta C/N_0}(t)$ is missing, a fourth-order polynomial fit of $\langle C/N_0(t)\rangle$ and $\sigma^2_{\Delta C/N_0}(t)$ over an entire satellite pass is adequate to obtain estimates of the noise variance. Otherwise, since a receiver implements the same

tracking loop across all channels, a noise estimate from another satellite may be applied. Fourth-order fits of $\langle C/N_0(t) \rangle$ and $\sigma^2_{\Delta C/N_0}(t)$ for all receivers are used for the analysis. The quality of the fits is demonstrated as the dashed lines in the top panels of Figure 6.8.

To prove the similarity of $\Delta C/N_0$ for each of the three receivers, the span of $\Delta C/N_0 \pm 2\sqrt{\sigma^2_{\Delta C/N_0}(t)}$ (i.e. the fade plus or minus two standard deviations of noise) for any one receiver must lie in the span of the other two receivers. If the measurement errors are gaussian, at least 95.45% of the measurements (two standard deviations) must lie within this span (or 4.55% will lie outside of the span).

Returning back to the PRN 17 data from the WAAS Houston (ZHU) Reference Station, the analysis is conducted. For the entire pass of the satellite (from rise to set), only 6.04% of carrier-to-noise measurements from each of the three receivers lie outside of the two-standard deviation span (the sample set consisted of 29706 one-second points). When constrained to times of fades greater than 1 dB during the solar radio burst (from 1800 - 2000 UT), 27.82% of the carrier-to-noise measurements from each of the three receivers lie outside the two-standard deviation span (sample size of 1319 one-second points). When receiver 1 is compared to receiver 2 over the entire pass, 1.71% of the points lie outside the two-standard deviation span, and during the solar radio burst only 8.64% of the points lie outside the span. When receiver 1 is compared to receiver 3, 3.08% of the points lie outside the span over the entire satellite pass, and 16.53% lie outside the span during the solar radio burst. Finally, when receiver 2 is compared to receiver 3, 2.60% of the points lie outside the span during the entire pass, and 14.25% lie outside the span during the solar radio burst. The statistics fall outside of the

expected ranges during the solar radio burst periods because each receiver within the Station operates with different tracking loop parameters and will respond dissimilarly to rapid changes in carrier-to-noise ratio. Receiver 1 and 2, however, have carrier-to-noise ratios that are statistically equivalent and receiver 3 is the outlier, at times. The peak fades observed at each receiver, however, are observed to be statistically equivalent, which is adequate motivation to determine the solar radio burst flux from the GPS measurements.

One-second carrier-to-noise ratio measurements from the WAAS Reference Stations are used to perform the analysis. Data from each of three receivers within each Reference Station for the satellite with the highest carrier-to-noise ratio at approximately 1940 UT on 5 December 2006 are averaged together. Carrier-to-noise ratios from the same satellite are used to create an average carrier-to-noise ratio for 6 December 2006 from all three receivers within the Reference Station. The satellite with the highest carrier-to-noise ratio at 1940 UT on 5 December 2006 is used because this time, shifted by a sidereal day, corresponds to 1936 UT on 6 December 2006 at the peak of the solar radio burst at L1. Furthermore, the highest carrier-to-noise ratio at 1940 UT on 5 December 2006 will correspond to the highest minimum value at the peak of the solar radio burst. In exactly the same way as $\Delta C/N_0$ was determined earlier, the average carrier-to-noise ratio from 6 December 2006 is subtracted from the average carrier-to-noise ratio from 5 December 2006, shifted by a sidereal day, to yield an estimate of the fade observed at each Reference Station.

Only data from the WAAS Reference Stations listed in Table 6.2 are included. Other Reference Stations are not included because 1-second carrier-to-noise ratios are not available, the Reference Station was not illuminated by the sun, or the

Table 6.2: The FAA WAAS reference stations used to determine the radiation pattern of the solar radio burst. The entries are sorted by solar elevation angle at 1936 UT.

Station	Location	Solar Elv.	Description
ZBW	42.7°N 71.5°W	12.7°	Nashua, NH
ZNY	40.8°N 73.1°W	15.0°	New York, NY
ZOB	41.3°N 82.2°W	18.7°	Cleveland, OH
ZAU	41.8°N 88.3°W	20.6°	Aurora, IL
BIL	45.8°N 108.5°W	21.3°	Billings, MT
ZSU	18.4°N 66.0°W	26.0°	San Juan, PR
ZTL	33.4°N 84.3°W	26.3°	Atlanta, GA
ZDV	40.2°N 105.1°W	26.4°	Denver, CO
ZME	35.1°N 90.0°W	27.2°	Memphis, TN
ZJX	30.7°N 81.9°W	27.3°	Jackson, MI
HNL	21.3°N 157.9°W	30.3°	Honolulu, HI
ZMA	25.8°N 90.3°W	30.3°	Miami, FL
ZAB	35.2°N 106.6°W	31.5°	Albuquerque, NM
ZLA	34.6°N 118.1°W	32.8°	Los Angeles, CA
ZHU	30.0°N 95.3°W	33.7°	Houston, TX

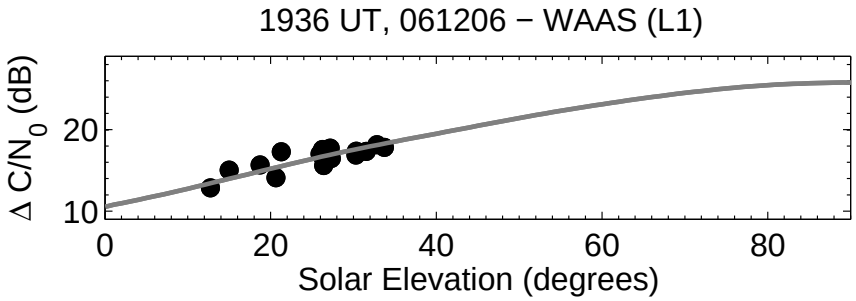

Figure 6.9: Plot illustrating the solar elevation angle at 1936 UT versus the peak fade measured at each of 15 WAAS Reference Stations. The gray line represents the estimated fade one should see on the WAAS network at all solar elevation angles, assuming a mean solar radio burst power of 926,000 SFU. The mean solar radio burst power is determined from the recorded fades on the 15 WAAS Reference Stations from which carrier-to-noise ratio data are available. For clarity error bars are omitted, but are shown in Figure 6.12

data were extreme outliers (e.g. the sun was behind a mountain). As expected, Reference Stations at lower solar elevation angles exhibit a smaller fade, as shown in Figure 6.9, for the largest fade observed on L1 at 1936 UT. $\Delta C/N_0$ for each Reference Station is near to the expected fade (gray line) assuming a peak solar radio burst power at L1 of approximately 500,000 SFU, as reported by OVSA. In Figure 6.10, three panels illustrate the magnitude of the fade at various times. Panel A shows the fades observed on the WAAS network before the onset of the solar radio burst, while Panel B shows a point during the fade, and Panel C shows the response of the network at the peak of the solar radio burst. The data points in Figure 6.9 correspond to Panel C. Note that in general the fades are more severe for the Stations located nearest to the subsolar point. A movie, depicting the fades observed on the WAAS Network as a function of time from approximately 1915 UT to 1945 UT is found on the accompanying CD (WAAS_061206_fade.mov).

Equation 4.7 determines the necessary solar radio burst flux to cause a 3 dB fade in received carrier-to-noise ratio. Although it may be used to obtain a rough idea of the magnitude of a fade to expect for larger solar radio bursts, an exact equation to convert any size fade in solar flux units is desired.

Working in decibels, a fade, $\Delta C/N_0$, is calculated as:

$$\Delta C/N_0 \;=\; C/N_{0_1} - C/N_{0_2} \text{ dB} \tag{6.16}$$

where C/N_{0_1} is the expected carrier-to-noise ratio without the presence of the solar radio burst, and C/N_{0_2} is the measured carrier-to-noise ratio in the presence of the solar radio burst. Recall that a positive measure of $\Delta C/N_0$ indicates a decrease in the received carrier-to-noise ratio as compared to the unaffected carrier-to-noise

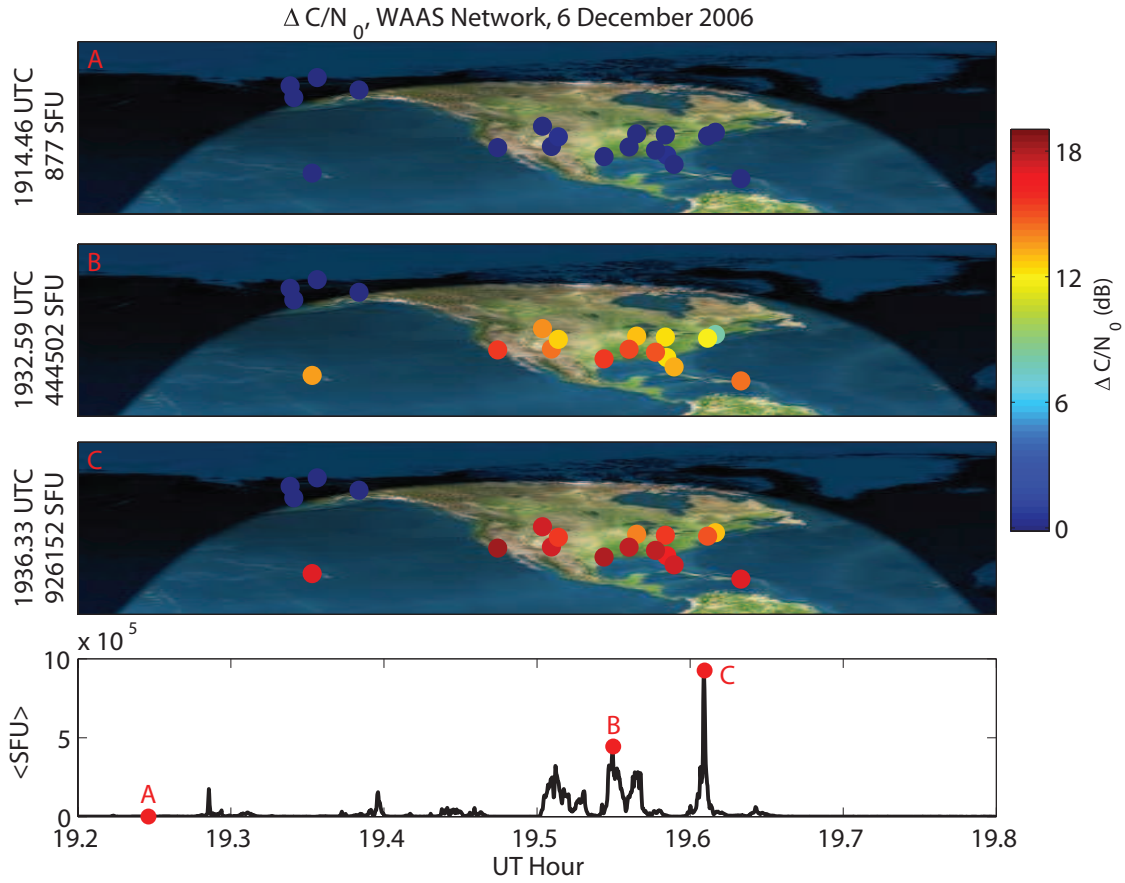

Figure 6.10: Measure of fading observed over the WAAS Reference Station network. Panel A illustrates that the measured fade prior to the solar radio burst is near zero. Panel B and Panel C illustrate the observed fades during the solar radio burst at particular times, noted to the left of each plot. Each colored dot represents the fade observed at that Reference Station. The bottom panel illustrates the mean solar radio burst RHCP power derived from the measured fades at each point in time. Note that in general, the Stations closest to the subsolar point exhibit a larger fade, as expected. The subsolar point lies due south of Mexico in the southern hemisphere. Image of Earth © The Living Earth, 1996 and is used here by permission of the publisher. Day/night overlay created using Earth Viewer by John Walker.

ratio. Setting $C/N_{0_1} = C/N_0$ allows C/N_{0_2} to be written in terms of C/N_0 by the addition of a term J, which is a broad-band noise interference source:

$$\Delta C/N_0 = \frac{C}{N_0} - \frac{C}{N_0 + J} \text{ dB} \tag{6.17}$$

Recall that N_0 is equivalent to the noise temperature of the receiver, P_n in $\frac{\text{W}}{\text{Hz}}$ from Equation 4.4. Moving from decibels to standard units, substituting $P_n = N_0$, and solving for J:

$$\begin{aligned} \Delta C/N_0 &= \frac{\frac{C}{P_n}}{\frac{C}{P_n+J}} \\ &= \frac{P_n + J}{P_n} \\ \Delta C/N_0 P_n &= P_n + J \\ P_n\left(\Delta C/N_0 - 1\right) &= J \text{ W/Hz} \end{aligned} \tag{6.18}$$

where J is noise from the solar radio burst. J is equivalent to the solar radio burst flux, P_{SRB}, in $\frac{\text{W}}{\text{m}^2-\text{Hz}}$ multiplied by the effective collecting area of the GPS antenna, A_{eff} in m^2 (Equation 4.5):

$$J = P_{SRB} A_{eff} \text{ W/Hz} \tag{6.19}$$

So the solar radio burst flux is measured as:

$$P_{srb} = \frac{P_n}{A_{eff}}\left(\Delta C/N_0 - 1\right) \ \frac{\text{W}}{\text{m}^2 - \text{Hz}} \tag{6.20}$$

Again, all quantities in Equation 6.20 are in standard (non-logarithmic) units.

The fade at each receiver is evaluated with Equation 6.20 to obtain the RHCP solar radio burst power at each sample time (nominally every second). At each sample time for each WAAS Reference Station, the solar elevation angle is calculated according to *Reda and Andreas* (2005) and determines the gain of the WAAS antenna at L1. A constant sky temperature of 100 K and receiver noise temperature of 2 dB are assumed. The NOAA reported temperature near-to the receiver location at 1930 UT is used for $T_{ambient}$ since the low-noise amplifier is located in the outdoor antenna. Finally, the solar radio burst power is adjusted for $\frac{1}{R^2}$ path loss (practically negligible) for comparison to flux measurements at the Owens Valley Solar Array.

The results are shown in Figure 6.11, and the accompanying movie (WAAS_061206_SFU.mov) is found on the enclosed CD. The estimated solar radio burst power at each Reference Station is not equal. The variance of the carrier-to-noise measurements into solar radio burst power explains some of the discrepancies resulting in biased C/N_0 measurements. Small differences in the antenna gain and noise figures from one Reference Station to another also alter the noise power spectral density of the receiver, resulting in biased C/N_0 measurements. These biases will clearly affect the estimate of the solar radio burst power.

A more detailed view at the peak of the fade, represented by Panel C in Figure 6.11, is shown in Figure 6.12. The top panel of this figure is the same as Figure 6.9, with the addition of error bars to indicate two standard deviations in $\Delta C/N_0$. At two standard deviations, most of the fades all lie within the expected fade given a mean RHCP solar radio burst power of 926,000 SFU. The bottom panel, for the same set of Reference Stations and time as the top panel, illustrates the RHCP

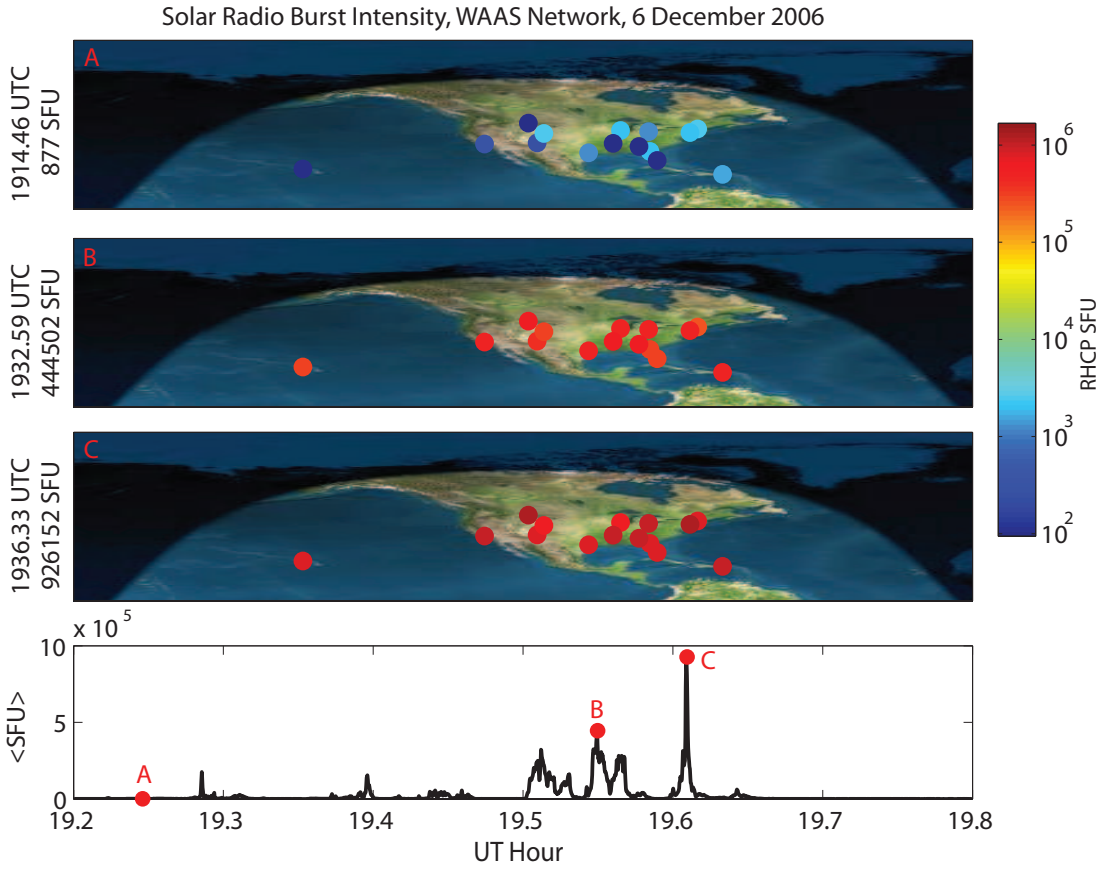

Figure 6.11: Estimated RHCP solar radio burst power derived from each WAAS Reference Station at various times (same as in Figure 6.10). The derived solar radio burst power is not as uniform as expected. The variance in the solar radio burst estimates is attributed to the variance in the carrier-to-noise ratio. Image of Earth © The Living Earth, 1996 and is used here by permission of the publisher. Day/night overlay created using Earth Viewer by John Walker.

solar radio burst estimate obtained from the fade. The error bars represent the estimated RHCP solar radio burst power given two standard deviations measured on $\Delta C/N_0$. The gray horizontal line represents the mean RHCP solar radio burst power measured at this time, which lies within the error bars for most of the Reference Stations.

Figure 6.13 illustrates how the WAAS Reference Station estimates of RHCP solar radio burst power compare to the values obtained from the Owens Valley

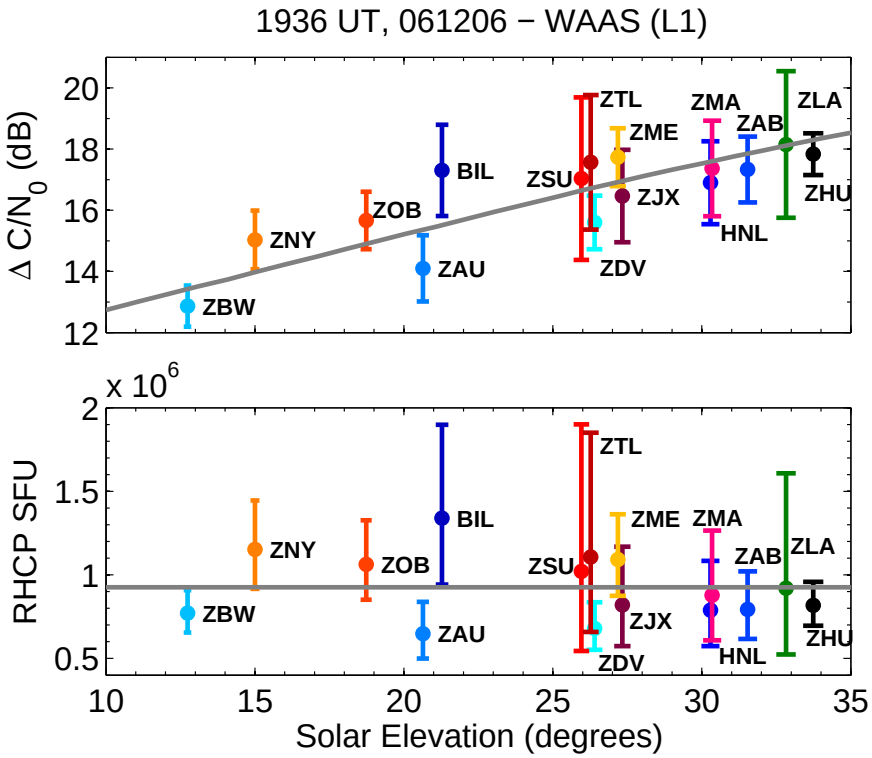

Figure 6.12: The top panel is the same as Figure 6.9, with the addition of error bars to indicate two standard deviations in $\Delta C/N_0$. The error bars of the fades lie within the expected fade, assuming a mean RHCP solar radio burst power of 926,000 SFU (gray line). The bottom panel illustrates how the standard deviation in $\Delta C/N_0$ corresponds to the possible range of RHCP solar radio burst estimates for each Reference Station. The estimated mean solar radio burst power at this time (gray line) lies within the error bars for all the Reference Stations. Both panels are for the same time (1936.38 UTC) as illustrated in Figure 6.11, Panel C.

Solar Array. Overall, the estimates from the WAAS Reference Stations agree well with the values measured at the Owens Valley Solar Array. At the peak of the fade, OVSA measured 504,000 SFU RHCP, whereas this technique estimates about 926,000 SFU. This is a large discrepancy, but is not alarming given the large spread of peak values measured even by dedicated solar radio burst equipment (e.g. Table 5.1). The peak flux measured using the WAAS receivers is bounded by the OVSA and the FST. Again, the discrepancies are expected given the different processing bandwidths, different operating frequencies, and different sample intervals

employed in WAAS receivers as compared to the OVSA.

Over the 1 hour span shown below, the RMS error is 22,000 SFU, with a peak error of 199,000 SFU occurring at 1936.38, at the peak solar radio burst intensity. The RMS error was calculated by averaging WAAS derived flux over the same 8 second time-interval as the OVSA data. It is believed that the true RMS error may be substantially lower because the OVSA saturated frequently during the bursts and reported erroneous values that will clearly skew the RMS error. The WAAS estimates, however, have some quirks in the data. In particular, the noise in $\Delta C/N_0$ makes it particularly difficult to measure small solar radio flux values. Solar flux estimates below 700 SFU are particularly noisy and difficult to detect due to the measurement noise in $\Delta C/N_0$. The strong similarity (zero-lag cross-correlation coefficient of 0.647) between the OVSA measurements and the GPS results suggests that GPS receivers could be used to monitor solar radio burst events.

It should be noted that other techniques were attempted to obtain estimates of the solar radio burst power. The three receivers within each Reference Station were used independently (as opposed to averaging), but the results were comparable. Likewise, smoothing the carrier-to-noise ratio for the prior day before computing $\Delta C/N_0$ did not significantly alter results. Finally, other methods (other than a simple average at each point in time) were attempted to obtain estimates across all the Reference Stations for the solar radio burst flux. These methods included a numerical least-squares estimate, a non-linear least-squares estimate, and a weighted non-linear least-squares estimate, with the weighting based on the variance of the carrier-to-noise ratios at each point in time. These methods resulted in estimates that are comparable to the average (within a few tens of thousands of SFU at the

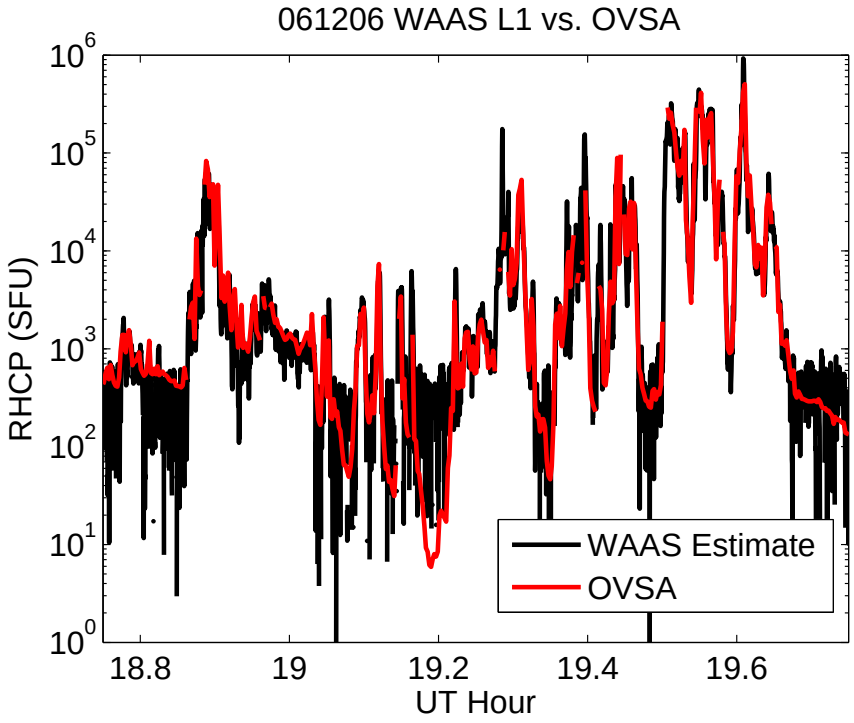

Figure 6.13: The WAAS Reference Station estimate of the RHCP solar radio burst power at GPS L1 is plotted against to the OVSA values at 1.6 GHz for the solar radio burst on 6 December 2006. The agreement between the WAAS estimate and the OVSA values is quite good. The WAAS estimates are particularly noisy below about 700 SFU due to the measurement noise in $\Delta C/N_0$.

peak fade) and are not included.

More extensive modeling techniques could be used to invert the measured fades in solar radio burst flux. For example, it is believed that the large discrepancies between measured carrier-to-noise ratios from one Reference Station to the next is in part due to varying noise figures. The author suggests that a relative calibration between GPS receivers could be made by comparing the received carrier-to-noise ratios from geostationary WAAS satellites.

6.4 Validity of Historical Record of Solar Radio Bursts

Three extra-ordinarily large solar radio bursts at 1 - 2 GHz over the span of 8 days is quite rare. As shown by *Nita et al.* (2002), solar radio bursts typically exhibit a power-law distribution of time between bursts (x-axis) versus peak intensity (y-axis), as shown in Figure 6.14. The records in the figure are generated from the National Geophysical Data Center (NGDC) of the National Oceanic and Atmospheric Administration for all solar radio bursts from 1960-1999. The data are supplied to various solar radio observatories around the world. Over the 40-year interval, observatories have been added and removed, the frequency of observation has changed, and the instrumentation has likely changed. As a result, the database may be biased due to incomplete observing coverage (missed events), and limitations in detection sensitivity at small fluxes and saturation of receivers at high fluxes.

The peak flux for each solar radio burst event in the database is recorded and logarithmically binned. For the solar maximum bins, 12,765 events were analyzed and 3464 events were analyzed for solar minimum. The power-law lines in the plot are drawn by eye and do not represent the actual fits determined by *Nita et al.* (2002), but are adequate for purposes of this discussion. The roll-off in solar radio bursts at flux densities below a few tens of solar flux units is attributable to instrument sensitivity. On the other hand, the roll-off at very high flux events may be attributable to a limitation in the scale size of solar flares as suggested by *Aschwanden et al.* (1998), or due to saturation effects of the instrumentation or missed observations due to geographical placement of the observatories.

The December 2006 solar radio bursts are indicated by the red dots on the line drawn for solar minimum, when these events occurred. Assuming these obser-

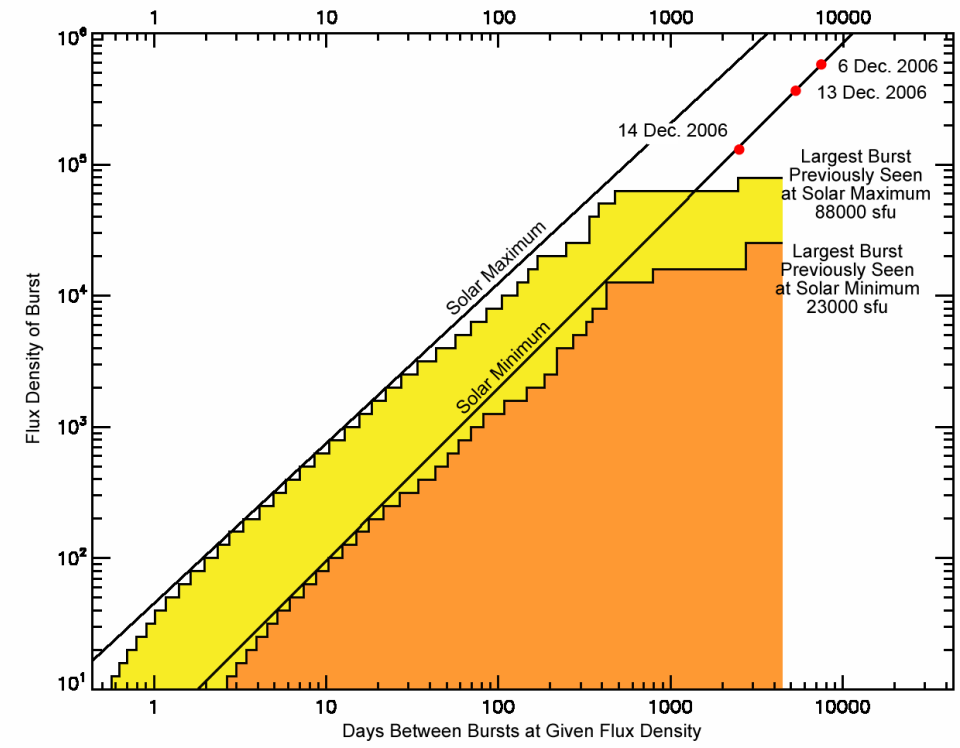

Figure 6.14: Histogram illustrating the frequency of solar radio bursts from 1 - 2 GHz at given intensities. The intensity of the solar radio bursts follows a power law. The December 2006 events are indicated by the red dot along the solar minimum line. The short time between and the large size of the December 2006 events suggests that the solar radio burst record is incomplete. The data are assembled by *Nita et al.* (2002). Figure courtesy of Dale Gary, NJIT.

vations are correct, the solar minimum line suggests that a solar radio burst the magnitude of the 6 December 2006 event (500,000 SFU by the OVSA estimate) should occur roughly every 8000 days, or roughly every 21 years. A solar radio burst the magnitude of the 13 December 2006 event should occur once every 4000 days, or 10 years. Finally, a solar radio burst the size of the 14 December 2006 event should occur every 2000 days, or every 5 years.

The size and occurrence of the December 2006 solar radio bursts suggests that the historical record may be inaccurate, and that limiting processes in solar flare physics may not be responsible for the upper limit in flux density. The events,

however, may just be a statistical fluke. Although it is beyond the scope of this thesis to explore this interesting phenomenon, the author does point out one telling fact: there appears to be a saturation level in the occurrence of solar radio burst power at about 100,000 SFU, which happens to be the saturation level of one particular instrument used in the solar radio burst record.

The United States Air Force (USAF) Radio Solar Telescope Network (RSTN) array is one contributor to the NGDC database on solar radio bursts. In the frequency range of 1 - 2 GHz, they report only solar radio burst total flux density at 1.415 GHz. The USAF RSTN documentation[9] indicates that, at 1.415 GHz, their instrument saturates at a level of 100,000 SFU total polarization.

For example, for 6 December 2006, the USAF RSTN station at Sagamore Hill, MA, reported a peak flux density at 1.415 GHz of 13,000 SFU at 1917 UT. The Palehua, HI station reported a peak flux density of 150,000 SFU at 1917 UT and again at 1931 UT. In comparison, the fades measured at the OVSA (RHCP only) at 1.4 GHz are 53,000 at 1917 UT and 341,000 at 1931 UT. The peak fade at OVSA is 1,003,000 SFU at 1933 UT. On 13 December 2006, the 1.415 GHz measurements at Palehua recorded a peak intensity of 150,000 SFU at 0225 UT, and at Learmonth, Australia, a peak intensity of 130,000 was recorded at 0336 UT and an event of 110,000 SFU was recorded at 0402 UT. According to the USAF RSTN record, there was also a 57,000 SFU burst recorded at 0437 UT on 13 December 2006. In contrast, the GPS L1 data suggest a peak intensity of 55,000 SFU at 0228 UT and a peak intensity of 184,000 SFU at 0230 on the L2 frequency. Finally, 14 December 2006 event registered in at 40,000 SFU on the Palehua Observatory at 2308 UT and at 63,000 SFU on the Learmonth Observatory. This event agrees well with

[9] ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_RADIO/RSTN_Radio_Solar_Telescope_Network_1second_data/RSTN.DOC

the flux estimates from the OVSA estimates at 1.4 GHz (48,000 SFU at 2307 UT), but not with the GPS estimate of 75,000 SFU at 1.6 GHz. Clearly, for the larger December 2006 solar radio bursts, the RSTN network is suffering from saturation effects.

The saturation of the USAF RSTN instruments during these events casts doubt on the accuracy of historic solar radio burst intensities. Did other observatories contributing to the NGDC database have similar saturation effects and at what level? Furthermore, missed observations due to the geographical location of the observatories might also contribute to missed events. GPS records are available from IGS and other sources for the past decade that could be used to reevaluate the historic record using the technique explored in Section 6.3.

6.5 Conclusion

Solar radio bursts challenge the ability of the receiver to track the satellite signals by significantly decreasing the carrier-to-noise ratio. Solar radio bursts are interfering wide-band (greater than 2 MHz for the C/A code) sources with relatively long duration (with respect to bit transitions). For this reason, traditional tactics of mitigating narrow-band interferers with adaptive filters or pulse-blanking are not practical. Instead, techniques to mitigate wide-band noise interference sources are presented. The techniques include reduction of noise bandwidth used in the carrier tracking loops, as well as minimization of clock noise and oscillator-induced vibration, utilization of bit prediction techniques, implementation of tightly-coupled IMUs, and increasing pre-correlation bandwidths and digitization levels. Null-steering antenna arrays provide good defense, and signal augmentations help to

mitigate interference problems. All the techniques require various degrees of modification to the receiver hardware and software, or augmentation to the actual satellite constellation.

A technique to determine solar radio burst flux density using carrier-to-noise ratio measurements from GPS receivers is also presented. Initial results indicate that the technique yields results that agree well with other instruments given differences in measurement frequencies, bandwidths, and sampling rates.

Finally, the validity of the historical record of solar radio bursts has been challenged by the December 2006 solar radio burst events. Several instruments underreported the magnitude of the bursts as they were plagued by saturation. Historical records for GPS span approximately 10 years, including the last solar maximum cycle, and could be used, in conjunction with the technique introduced in this chapter, to determine if other highly energetic solar radio burst events in the GPS frequencies were not accurately reported.

Acknowledgements

Dale Gary of the New Jersey Institute of Technology provided both the OVSA and FST data for the December 2006 solar radio bursts and results of the detailed analysis of the 6 December 2006 event. Gratitude is also extended to Pat Reddan of Zeta Associates Inc. for providing valuable resources on WAAS, including antenna gain patterns and references to papers. 1-second WAAS data was provided by Tom McHugh of the FAA and also by Patricia Doherty of Boston College. The views reported here are solely the views of the author and not of the FAA or other government agencies.

Chapter 7

Conclusion

We've arranged a civilization in which most crucial elements profoundly depend on science and technology. - Carl Sagan

This thesis has shown qualitatively and quantitatively how a solar radio burst affects GPS receivers. Each chapter has its own conclusion, so it is not appropriate to reiterate what has already been said. Rather, some broader impacts of the December solar radio bursts will be highlighted to stress how unawares many institutions and individuals were caught by the severe space weather events.

In a NOAA memorandum concerning the space weather impacts and activity from NOAA region 10930[1], the entire sunspot group (AR 10930) seems to have surprised the forecasters at the Space Environment Center (now Space Weather Prediction Center):

The strong flare activity on December 5 was unanticipated. Region 930 had not given SEC forecasters any reason to think it would be such a prolific flare producer as it was just then rotating into view. Its R3-level flares belied the fact that December 2006 was almost 7 years post solar maximum, and climatology would predict X class flares to occur on less than 1% of the days near solar minimum.

After the initial X8 flare on 5 December 2006, NOAA continued to monitor the sunspot and successfully warned of future space weather impacts from the sunspot group. The solar radio burst was noted by the Australian government and a

[1]Portions of this chapter are based on the publication *Cerruti et al.* (2007), which is joint collaborative work with A. P. Cerruti, P. M. Kintner, D. E. Gary, A. J. Mannucci, R. F. Meyer, P. Doherty, and A. J. Coster. Reproduced by permission of the American Geophysical Union.

[1]The memorandum was written for for Brigadier General David L. Johnson, Assistant Administrator for Weather Services and Director, National Weather Service

warning was issued[2]:

The region responsible for this solar activity is still visible on the suns disk and there is the strong possibility of further activity and GPS failures before the rotation of the sun takes the active region out of view.

The NOAA memorandum also cites a report from GPS Operations Center at Schriever Air Force Base that noted:

"At approximately 6 Dec/2000Z there was a widespread loss of GPS in the Mountain States region, specifically around the 4 corners region of NM/CO. Several aircraft reported losing lock on GPS and were tracking 7-9 satellites, and abruptly loss [sic] *locks and were then tracking 0-1."*

The statement from GPS Operations Center is noteworthy because they only acknowledge service disruptions to GPS of significant duration. It is unclear whether the aircraft were military, or if they were utilizing encrypted military receivers. A recent letter to *GPS World* from the Space and Missile Systems Center's GPS Wing at Los Angeles Air Force Base claims that there is no evidence of an impact on military GPS receivers. They claim that military users continued to receive uninterrupted service and that the NOAA memorandum was a reference to civil users[3]. It is unclear to the author that military receivers, which use the encryption key to track P(Y) signals on L1 and L2, should be more robust than civilian receivers. The wide-band nature of the solar radio burst will be equally detrimental to carrier-to-noise ratio. Military receivers likely incorporate inertial measurement units as well as interference mitigation technologies that would make them less prone to solar radio bursts. There is a possibility that the solar radio

[2]http://www.ips.gov.au/Category/Site%20News/media-release-plus-updateDec2006.pdf
[3]http://sidt.gpsworld.com/gpssidt/article/articleDetail.jsp?id=444832

burst may not have affected receivers to the point that the Precise Positioning Service mandates were exceeded. The Precise Positioning Service metric is not published, but (*Cox Jr.*, 1994) suggests accuracy of at least 16 meters horizontally and 10 meters vertically.

The memorandum also appears to indicate that the solar radio burst also appears to have affected air traffic control radar systems at Denver International Airport. These reports are not without precedent as the impact of solar radio bursts on radar systems have been noted by others (e.g. *Kudeki et al.* (2006) and *Scro and Quigley* (2006)). The impact of solar radio bursts on other radio technologies such as cell-phone networks is also well documented (*Lanzerotti et al.*, 1999).

The December 2006 solar radio burst impacts on GPS were noted on other systems as well. *Cerruti et al.* (2007) indicate that the NASA Global Differential GPS (GDGPS) system, which is a real-time differential GPS system that provides decimeter positioning accuracy and sub-nanosecond timing accuracy. The GDGPS corrections for satellites within the SRB affected service volume (satellites only visible from the sunlit hemisphere) were unavailable for several minutes, and caused degraded or unavailable positioning accuracy for some users. Although the effects of the solar radio burst were noted within one minute, the service was affected for over half an hour on 6 December.

Cerruti et al. (2007) indicate that the most serious impacts will be on those users requiring truly continuous service. *Cerruti et al.* (2007) give in particular one example:

"One example are off-shore oil platforms that rely on precision corrections to

maintain their position. Fugro Chance, Inc., a provider of precision GPS ranging corrections, reported that if the solar radio burst on 6 December 2006 had lasted several minutes longer, it would have had a severe impact on its customers, including off-shore oil platforms (Richard Barker, Space Weather Enterprise Forum, Washington, DC, 2007). From Fugro Chance's viewpoint, a major need is the ability to predict space weather, providing operational flexibility."

As shown in this thesis, the impact of solar radio bursts on users of GPS can vary from a minor interruption of service and data to major operational and financial impacts. These impacts depend on the solar radio burst intensity and duration and the technology of the receivers. Understanding the future impact of solar radio bursts requires intimate knowledge of GPS technologies and how they are used. Future impacts of solar radio bursts also depends on the frequency and intensity of future radio bursts.

The frequency and intensity of future solar radio bursts is most difficult to predict, because, as shown in Chapter 6, the historical record appears to be incorrect. If the historical record of very intense solar radio burst observations has been compromised, then additional events of an intensity that can impact GPS during the next solar maximum in 2011 and the 3-4 surrounding years are possible. *Cerruti et al.* (2007) indicate that if the historical record is accurate, then an explanation for the intense solar radio bursts of December 2006 is required and currently no such explanation exists.

At Space Weather Week 2007 held in Boulder, Colorado, the author held a conversation with a Navy individual about an incident involving loss of GPS service for aircraft in the South Pacific Ocean in the early 1990s. Although the individual could never prove the ultimate cause of the event, he is fairly certain that the

culprit was a solar radio burst. Again, if that is the case, it would appear that the historical record is incorrect. At the same conference, the inadequacy of the USAF RSTN network to properly measure solar radio burst flux density seemed to have at least spurred another individual from the Air Force Research Lab to invest money for improving the network's capability.

Given the uncertainty in the historical record of solar radio bursts coupled with the broad impact of the solar radio burst leads to uncertainty about what to expect and how to prepare for the next solar maximum period and beyond. Ultra-precise positioning and timing services have become overly reliant on a system that has consistently exceeded the performance standards set in *GPS SPS Performance Standard* (2001). While no where does the United States Government guarantee truly continuous and accurate (< 5 m) services with GPS, the current capabilities of the system has launched business plans and expectations that naively assume the current level of service is always available. To paraphrase Carl Sagan's quote at the top of the page, we have become overly reliant on GPS and other wireless technologies without an awareness and respect for the potential effects of future space weather events.

APPENDIX

A.1 Dynamic Stress Error on Tracking Loops

Dynamic stress error is the error induced by satellite-receiver relative motion. The order of the tracking loop determines the type of stress it is sensitive to. For a PLL, a second-order loop under acceleration will suffer a steady state error. A third-order PLL is sensitive to jerk. From *Ward* (1996b), the dynamic stress error for a second-order loop with minimum mean square error is given as:

$$\theta_{e2} = \frac{dR^2/dt^2}{\left(\frac{B_L}{0.52}\right)^2} = 0.2809\frac{dR^2/dt^2}{B_L^2} \text{ degrees} \tag{A.1}$$

where B_L is the loop noise bandwidth, and dR^2/dt^2 is the maximum line-of-sight acceleration dynamics in deg/sec^2.

For a third order loop with minimum mean square error, the dynamic stress is:

$$\theta_{e3} = \frac{dR^3/dt^3}{\left(\frac{B_L}{0.7845}\right)^3} = 0.4828\frac{dR^3/dt^3}{B_L^3} \text{ degrees} \tag{A.2}$$

where dR^3/dt^3 is the maximum line-of-sight jerk dynamics in deg/sec^3.

An FLL tracking loop contains one more integrator than the PLL tracking loop of the same order, so the stress error for an n$^{\text{th}}$ order FLL is:

$$f_e = \frac{d}{dt}\left(\frac{1}{360\frac{B_L^n}{F}}\frac{dR^n}{dt^n}\right) \tag{A.3}$$

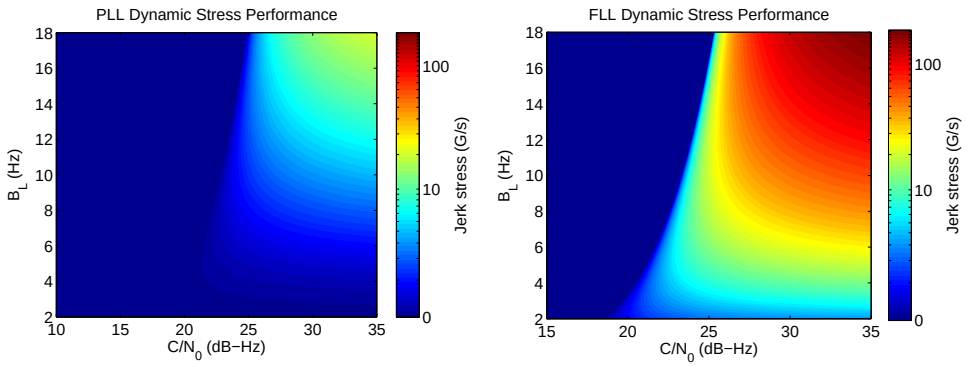

Figure A.1: The left panel illustrates the maximum jerk stress (G/s) tolerated by a third-order PLL at a given carrier-to-noise ratio, C/N_0, and noise loop bandwidth, B_L. The right panel shows the maximum jerk stress for a second-order FLL, which is equivalent to the third-order PLL. In both cases, the pre-correlation integration time is $T = 10$ ms.

where $F = 0.25$ for n = 1, and $F = 0.53$ for n = 1.

Figure A.1 illustrates the maximum jerk stress tolerated by a third order PLL and an equivalent third order FLL as a function of noise bandwidth and carrier-to-noise ratio, assuming only thermal noise and dynamic stress contribute to the tracking error. In each case, the pre-correlation integration time, T, is equal to 10 ms. Note that the second-order FLL has roughly a factor of ten better stress performance than the third-order PLL. For both the PLL and FLL, decreasing the pre-correlation integration time results in more tolerance of jerk-stress at a given carrier-to-noise, however a penalty is pain in that the minimum carrier-to-noise ratio at which the receiver can maintain lock also increases.

A.2 Receiver Vibration

In some cases, the receiver is in an environment where vibration is so severe that the effect of vibration on the oscillator must be considered. Minimizing vibration

is important because it helps to minimize the overall PLL phase error. When constructing a receiver, for example, to be utilized in construction equipment or on rockets, the receiver should be mounted on vibration isolators to function properly and maintain lock.

Vibration-induced oscillator jitter is given by (*Ward*, 1996b):

$$\sigma_v = \frac{360 f_L}{2\pi} \sqrt{\int_{f_{min}}^{f_{max}} S_v^2(f_m) \frac{P(f_m)}{f_m^2} df_m} \text{ degrees} \tag{A.4}$$

where f_L is the operating frequency (L1 or L2) in Hz, f_m is the random vibration modulation frequency in Hz, $S_v(f_m)$ is the oscillator vibration sensitivity of $\delta f / f_L$ per G as a function of f_m, and $P(f_m)$ is the power curve of the random vibration in G^2/Hz as a function of f_m.

BIBLIOGRAPHY

Aarons, J. (1982), Global morphology of ionospheric scintillations, *Proc. IEEE*, *70*, 360–378.

Aarons, J. (1997), Global Positioning System phase fluctuations at auroral latitudes, *J. Geophys. Res.*, *102*(17), 17,219–17,231.

Allan, D. W. (1987), Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators, *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, *34*, 647–654.

Allan, D. W., N. Ashby, and C. Hodge (1998), Time in the space age, *IEEE Spectrum*, *35*, 42–51, doi:10.1109/MSPEC.1998.663757.

Aparicio, M., P. Brodie, L. Doyle, J. Rajan, and P. Torrione (1996), *Global Positioning System: Theory and Applications*, vol. 1, chap. 6: GPS Satellite and Payload, pp. 209 – 244, American Institute of Aeronautics and Astronautics, Washington, D.C.

Aschwanden, M. J., B. R. Dennis, and A. O. Benz (1998), Logistic avalanche processes, elementary time structures, and frequency distributions in solar flares, *Astrophys. J.*, *497*.

Barbour, N. M., J. M. Elwell, and R. H. Sutterlund (1992), Inertial Instruments: Where to Now?, in *AIAA Guidance, Navigation, and Control Conference*, pp. 566–574, American Institute of Aeronautics and Astronautics.

Barker, C., J. W. Betz, J. E. Clark, J. T. Correia, J. T. Gillis, S. Lazar, L. Rehborn, and J. R. Straton III (2000), Overview of the GPS M code signal, in *Proceedings of the 2000 National Technical Meeting of the Institute of Navigation*, pp. 542 – 549, The Institute of Navigation.

Bassiri, S., and G. Hajj (1993), Higher-order ionospheric effects on the global positioning system observables and means of modeling them, *J. Geod.*, *18(6)*, 280–289.

Bastian, T. S., A. O. Benz, and D. E. Gary (1998), Radio emission from solar flares, *Annu. Rev. Astron. Astrophys.*, *36*, 131–188.

Basu, S., E. J. Weber, T. W. Bullett, M. J. Keskinen, E. MacKenzie, P. Doherty, R. Sheehan, H. Kuenzler, P. Ning, and J. Bongiolatti (1998), Characteristics of plasma structuring in the cusp/cleft region at Svalbard, *Radio Sci.*, *33*, 1885 – 1900.

Basu, S., S. Basu, K. M. Groves, E. MacKenzie, M. J. Keskinen, and F. J. Rich (2005a), Near-simultaneous plasma structuring in the midlatitude and equatorial ionosphere during magnetic superstorms, *Geophys. Res. Lett.*, *32*, doi:10.1029/2004GL021678.

Basu, S., S. Basu, J. J. Makela, R. E. Sheehan, E. MacKenzie, P. Doherty, J. W. Wright, M. J. Keskinen, D. Pallamraju, L. J. Paxton, and F. T. Berkey (2005b), Two components of ionospheric plasma structuring at midlatitudes observed during the large magnetic storms of October 30, 2003, *Geophys. Res. Lett.*, *32*, doi:10.1029/2004GL021669.

Bauregger, F. (2003), Novel anti-jam antennas for airborne GPS navigation, Ph.D. thesis, Stanford University.

Beach, T. L., and P. M. Kintner (2001), Development and use of a GPS ionospheric scintillation monitor, *IEEE Trans. Geosci. Remote Sensing*, *39*, 918–928.

Betz, J. W. (2002), Binary offset carrier modulations for radionavigation, *Journal of the Institute of Navigation*, *48*(4), 227–246.

Cerruti, A. P., P. M. Kintner, D. E. Gary, L. J. Lanzerotti, E. R. de Paula, and H. B. Vo (2006), Observed solar radio burst effects on GPS/Wide Area Augmentation System carrier-to-noise ratio, *Space Weather*, *4*, doi:10.1029/2006SW000254.

Cerruti, A. P., P. M. Kintner, D. E. Gary, L. J. Lanzerotti, E. R. de Paula, and H. B. Vo (2007), The effect of intense December 2006 solar radio bursts on GPS receivers, *Space Weather*, in review.

Chen, F. F. (1984), *Introduction to plasma physics and controlled fusion*, vol. 1, chap. 4: Waves in Plasmas, pp. 79 – 154, Plenum Press, New York and London.

Chen, Z., Y. Gao, and Z. Liu (2005), Evaluation of solar radio bursts' effect on GPS receiver signal tracking within International GPS Service Network, *Radio Sci.*, doi:10.1029/2004RS003066.

Coker, C., G. S. Bust, R. A. Doe, and T. L. Gaussiran II (2004), High-latitude plasma structure and scintillation, *Radio Sci.*, *39*(RS1S15), doi:10.1029/2002RS002833.

Coster, A., S. Skone, C. Mitchell, G. D. Franceschi, L. Alfonso, and V. Romano (2005), Global studies of GPS scintillation, in *Proceedings of the 2005 National Technical Meeting of the Institute of Navigation*, Institute of Navigation.

Cox Jr., T. M. (1994), *PPS GPS: What is it? And how do I get it?*, U.S. Army Topographic Engineering Center.

Csillaghy, A., and A. O. Benz (1993), The bandwidth of millisecond radio spikes in solar flares, *Astronomy and Astrophysics*, *274*.

Czopek, F., and L. S. Shollenberger (1993), Description and performance of the

GPS Block I and II L-Band antenna and link budget, in *6th International Technical Meeting*, The Institute of Navigation.

Davies, K. (1965), *Ionospheric Radio Propagation*, National Bureau of Standards, Washington DC.

Dessler, A. J., and E. N. Parker (1959), Hydromagnetic theory of magnetic storms, *J. Geophys. Res.*, *64*.

Doherty, P., A. J. Coster, and W. Murtagh (2003), Space weather effects of October November 2003, *GPS Solutions*, *8*(4), doi:10.1007/s10291-004-0109-3.

Dow, J. M., R. E. Neilan, and G. Gendt (2005), The International GPS Service (IGS): Celebrating the 10th anniversary and looking to the next decade, *Adv. Space Res.*, *36*(3), 320–326, doi:10.1016/j.asr.2005.05.125.

Dungey, J. W. (1961), Interplanetary magnetic field and the auroral zones, *Phys. Res. Lett.*, *6*, 47.

Forbes, T. G., and L. W. Acton (1996), Reconnection and field line shrinkage in solar flares, *Astrophys. J.*, *459*, 330, doi:10.1086/176896.

Foster, J. C., and F. J. Rich (1997), Prompt midlatitude electric field effects during severe geomagnetic storms, *Geophys. Res. Lett.*, *103*, doi:10.1029/97JA03057.

Foster, J. C., and W. Rideout (2005), Midlatitude TEC enhancements during the October 2003 superstorm, *Geophys. Res. Lett.*, *32*(L12S04), doi:10.1029/2004GL021719.

Foster, J. C., W. Rideout, B. Sandel, W. T. Forrester, and F. J. Rich (2007), On the relationship of SAPS to storm-enhanced density, *J. of Atmosph. and Solar-Terrestrial Phys.*, *69*, doi:10.1016/j.jastp.2006.07.021.

Francisco, S. G. (1996), *Global Positioning System: Theory and Applications*, vol. 1, chap. 10: GPS Operational Control Segment, pp. 435–466, American Institute of Aeronautics and Astronautics, Washington, D.C.

Gao, G. X., A. Chen, S. Lo, D. D. Lorenzo, and P. Enge (2007), GNSS over China - The Compass MEO satellite codes, *Inside GNSS*, pp. 36–43.

GPS SPS Performance Standard (2001), *Global Positioning System Standard Positioning Service Performance Standard*, Department of Defense, http://gps.afspc.af.mil/gpsoc/documents/GPS_Signal_Spec.pdf.

Greenspan, M. E., C. E. Rasmussen, W. J. Burke, and M. A. Abdu (1991), Equatorial density depletions observed at 840 km during the great magnetic storm of March 1989, *J. Geophys. Res.*, *96*(13).

Hoque, M. M., and N. Jakowski (2007), Higher order ionospheric effects in precise GNSS positioning, *J. Geod.*, *81*, 259–268, doi:10.1007/s00190-006-0106-0.

Humphreys, T. E., M. L. Psiaki, J. P. M. Kintner, and B. M. Ledvina (2005), GPS carrier tracking loop performance in the presence of ionospheric scintillations, in *Proceedings of the 18th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2005*, pp. 156–167, The Institute of Navigation.

Humphreys, T. E., M. L. Psiaki, P. M. Kintner, and B. M. Ledvina (2006), GNSS receiver implementation on a DSP: status, challenges, and prospects, in *Proceedings of the 19th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2006*, pp. 2370–2382, The Institute of Navigation.

Humphreys, T. E., M. L. Psiaki, B. M. Ledvina, A. P. Cerruti, and J. P. M. Kintner (2007), Characterization of severe ionospheric scintillation and its effect on GPS carrier phase tracking, *IEEE Transactions on Aerospace and Electronic Systems*, in Review.

Hysell, D. L., and E. Kudeki (2004), Collisional shear instability in the equatorial *F* region ionosphere, *J. Geophys. Res.*

IS-GPS-200D (2004), *NAVSTAR Global Positioning System Interface Specification, IS-GPS-200 Revision D*, NAVSTAR GPS Joint Program Office.

IS-GPS-705 (2005), *Navstar GPS Space Segment / User Segment L5 InterfacesIS-GPS-200 Revision D*, NAVSTAR GPS Joint Program Office.

IS-GPS-800 (2007), *Navstar GPS Space Segment / User Segment L1C Interfaces IS-GPS-800*, Depratment of the Air Force Space and Missile Systems Center.

Kaplan, E. D. (1996), *Understanding GPS Principles and Applications*, chap. 1: Introduction, pp. 1 – 14, Artech House Publishers, Boston.

Kasper, J. F., and C. E. Hutchinson (1978), The OMEGA navigation system - an overview, *IEEE Communications Society Magazine*, pp. 23–35.

Kelley, M. C. (1989), *The Earth's ionosphere: Plasma physics and electrodynamics*, Academic Press, San Diego.

Kelley, M. C., J. J. Makela, B. M. Ledvina, , and P. Kintner (2002), Observations of equatorial spread-F from Haleakala, Hawaii, *Geophys. Res. Lett.*, *29*(20), doi: 10.1029/2002GL015509.

Kelley, M. C., J. J. Makela, and O. de la Beaujardiére (2006), Convective Iono-

spheric Storms: A Major Space Weather Problem, *Space Weather*, *4*, doi: 10.1029/2005SW000144.

Keskinen, M. J., S. Basu, and S. Basu (2004), Midlatitude sub-auroral ionospheric small scale structure during a magnetic storm, *Geophys. Res. Lett.*, *31*, doi: 10.1029/2003GL019368.

Kintner, P. M., H. Kil, C. Deehr, and P. Schuck (2002), Simultaneous total electron content and all-sky camera measurements of an auroral arc, *J. Geophys. Res.*, *107*(A7), 1127, doi:10.1029/2001JA000110.

Kintner, P. M., B. M. Ledvina, E. R. de Paula, and I. J. Kantor (2004), Size, shape, orientation, speed, and duration of GPS equatorial anomaly scintillations, *Radio Sci.*, *39*, doi:10.1029/2003RS002878.

Kintner, P. M., B. M. Ledvina, and E. R. de Paula (2007), GPS and ionospheric scintillations, *Space Weather*, *5*, doi:10.1029/2006SW000260.

Kivelson, M. G., and C. T. Russell (Eds.) (1995), *Introduction to space physics*, Cambridge Univ. Press, New York, Cambridge, and Melbourne.

Klobuchar, J. A., J. M. Kunches, and A. J. Van Dierendonck (1999), Eye on the ionosphere: Potential solar radio burst effects on GPS signal to noise, *GPS Sol.*, *3*, 69–71.

Kudeki, E., R. Woodman, P. Reyes, and J. L. Chau (2006), Ionospheric observations at Jicamarca during the Sept. 7, 2005 solar flare, in *CEDAR Workshop Proceedings*, CEDAR.

Kundu, M. R. (1965), *Solar Radio Astronomy*, John Wiley and Sons, New York.

Lanzerotti, L. J., D. J. Thomson, and C. G. Maclennan (1999), *Modern Radio Science 1999*, 25 - 50 pp., Wiley-IEEE Press.

Lavrakas, J. W. (2006), Managing the GPS Constellation for Today's Needs, *GPS World*.

Ledvina, B., J. J. Makela, and P. M. Kintner (2002), First observations of intense GPS L1 amplitude scintillations at midlatitude, *Geophys. Res. Lett.*, *29*(14), doi:10.1029/2002GL014770.

Ledvina, B. M., A. P. Cerruti, M. L. Psiaki, S. P. Powell, and P. M. Kintner (2003a), Performance tests of a real-time 12-channel GPS L1 software receiver, in *Proceedings of the 16th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GPS/GNSS 2003*, pp. 678–688, The Institute of Navigation.

Ledvina, B. M., M. L. Psiaki, and P. M. Kintner (2003b), A 12-channel real-time GPS L1 software receiver, in *Proceedings of the 2003 National Technical Meeting of the Institute of Navigation*, pp. 767–782, The Institute of Navigation.

Ledvina, B. M., M. L. Psiaki, S. P. Powell, and P. M. Kintner (2004a), Bit-wsie parallel algorithms for efficient software correlation applied to a GPS software receiver, *IEEE Transactions on Wireless Communications*, *3(5)*.

Ledvina, B. M., M. L. Psiaki, D. J. Sheinfeld, A. P. Cerruti, S. P. Powell, and P. M. Kintner (2004b), A real-time GPS civilian L1/L2 software receiver, in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2004*, pp. 986–1005, The Institute of Navigation.

Ledvina, B. M., M. L. Psiaki, T. E. Humphreys, S. P. Powell, and P. M. Kintner (2005), Real-time software receiver tracking of GPS L2 Civilian signals using a hardware simulator, in *Proceedings of the 18th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2005*, pp. 1598–1610, The Institute of Navigation.

Ledvina, B. M., M. L. Psiaki, T. E. Humphreys, S. P. Powell, and P. M. Kintner (2006), A real-time software receiver for the GPS and Galileo L1 signals, in *Proceedings of the 19th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2006*, pp. 2321–2333, The Institute of Navigation.

Leva, J. L., M. U. de Haag, and K. Van Dyke (1996), *Understanding GPS Principles and Applications*, chap. 7: Performance of Standalone GPS, pp. 1 – 14, Artech House Publishers, Boston.

Liu, Z., D. E. Gary, G. M. Nita, S. M. White, and G. J. Hurford (2007), A subsystem test bed for the Frequency-Agile Solar Radiotelescope, *Publ. Astronomical Society of the Pacific*, *119*, 303.

Makela, J. J., M. C. Kelley, J. J. Sojka, X. P, and A. J. Mannucci (2001), GPS normalization and preliminary modeling results of total electron content during a midlatitude space weather event, *Radio Sci.*, *36*(2).

Makela, J. J., B. M. Ledvina, M. C. Kelley, and P. Kintner (2004), Analysis of the seasonal variations of equatorial plasma bubble occurrence observed from Haleakala, Hawaii, *Annal. Geophys.*, *22*(9).

Mannuci, A. J., B. D. Wilson, D. N. Yuan, C. H. Ho, U. J. Lindqwister, and T. F.

Runge (1998), A global mapping technique for GPS-derive ionospheric total electron content measurements, *Radio Sci.*, *33*(3), doi:10.1029/97RS02707.

Manz, A. M., K. Shallberg, and P. Shloss (2000), Improving WAAS Receiver Rdaio Frequency Interference Rejection, in *Proceedings of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GPS 2001*, pp. 471–479, The Institute of Navigation.

Mendillo, M. (2006), Storms in the Ionosphere: Patterns and Processes for Total Electron Content, *Rev. of Geophys.*, *47*(9), doi:10.1029/2005RG000193.

Mishin, E. V., W. J. Burke, S. Basu, S. Basu, P. M. Kintner, and B. Ledvina (2003a), Stormtime ionospheric irregularities in SPAS-related troughs: Causes of GPS scintillations at midlatitudes, in *Fall Meet. Suppl.*, vol. 84, Eos Trans. AGU, abstract SH52A-07.

Mishin, E. V., W. J. Burke, C. Y. Huang, and F. J. Rich (2003b), Electromagnetic wave structures within subauroral polarization streams, *J. Geophys. Res.*, *108*(A8), doi:10.1029/2002JA009793.

Misra, P., and P. Enge (2006), *Global Positioning System, Signals Measurements, and Performance*, 2nd ed., Ganga-Jamuna Press, Lincoln, Massachusetts.

Mitchell, C. N., L. Alfonsi, G. D. Franceschi, M. Lester, and V. R. andand A. W. Wernik (2004), GPS TEC and scintillation measurements from the polar ionosphere during the October 2003 storm, *Geophys. Res. Lett.*, *32*(L12S03), doi:10.1029/2004GL021644.

Mitra, A. P. (1974), Ionospheric effects of solar flares, *Astrophysics and Space Science Library*, *46*.

Morrissey, T. N., K. W. Shallberg, A. J. Van Dierendonck, T. Kim, and M. J. Nicholson (2002), GPS receiver performance characterization under simulated ionospheric scintillation environments II, in *Proceedings of the 2002 National Technical Meeting of the Institute of Navigation*, pp. 682–693, The Institute of Navigation.

Nicolls, M. J., and M. C. Kelley (2005), Strong evidence for gravity wave seeding of an ionospheric plasma instability, *Geophys. Res. Lett.*, *35*, doi: 10.1029/2004GL020737.

Nita, G. M., D. E. Gary, L. J. Lanzerotti, and D. J. Thomson (2002), The peak flux distribution of solar radio bursts, *Astrophys. J.*, *570*, 423–438.

Nita, G. M., D. E. Gary, and J. Lee (2004), Statistical study of two years of solar flare radio spectra obtained with the Owens Valley Solar Array, *Astrophys. J.*, *605*, 528–545.

Parkinson, B. W. (1996a), *Global Positioning System: Theory and Applications*, vol. 1, chap. 1: Introduction and Heritage of NAVSTAR, the Global Positioning System, pp. 3–28, American Institute of Aeronautics and Astronautics, Washington, D.C.

Parkinson, B. W. (1996b), *Global Positioning System: Theory and Applications*, vol. 2, chap. 1: Differential GPS, pp. 3–50, American Institute of Aeronautics and Astronautics, Washington, D.C.

Psiaki, M., and S. Mohiuddin (2005), Modeling, Analysis and Simulation of GPS Carrier Phase for Spacecraft Relative Navigation, in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, American Institute of Aeronautics and Astronautics.

Reda, I., and A. Andreas (2005), *Solar Position Algorithm for Solar Radiation Applications*, National Renewable Energy Laboratory, http://www.nrel.gov/docs/fy06osti/34302.pdf.

Rideout, W., and A. Coster (2006), Automated GPS processing for global total electron content data, *GPS Solutions*, *10*(3), 219–228.

Rodrigues, F. S. (2003), Estudo das irregularidades ionoféricas equatoriais utilizando sinais GPS, Master's thesis, Instituto Nacional de Pesquisas Espaciais.

Rounds, S. (2004), Jamming Protection of GPS Receivers, Part II: Antenna Enhancements, *GPS World*.

Rushanan, J. J. (2007), The spreading and overlay codes for the L1C signal, in *Proceedings of the 2007 National Technical Meeting of the Institute of Navigation*, pp. 539–547, Institute of Navigation.

Sarwate, D. V., and M. B. Pursley (1980), Cross-correlation properties of pseudorandom and related sequences, *Proc. IEEE*, *68*(5), 593–618.

Schunk, R. W., and A. F. Nagy (2000a), *Ionospheres: physics, plasma physics, and chemistry*, chap. 11: The Terrestrial Ionosphere at Middle and Low Latitudes, pp. 312 – 365, Cambridge Univ. Press, Cambridge.

Schunk, R. W., and A. F. Nagy (2000b), *Ionospheres: physics, plasma physics, and chemistry*, Cambridge Univ. Press, Cambridge.

Sckopke, N. (1966), A general relation between the energy of trapped particles and the disturbance field near the Earth, *J. Geophys. Res.*, *71*.

Scro, K. D., and S. Quigley (2006), Operationalizing space weather products -

process and issues, in *AGU Fall 2006 Meeting*, American Geophysical Union, abstract SA51A-01.

Simon, M. K. (1978), On the Calculation of Squaring Loss in Costas Loops with Arbitrary Arm Filters, *IEEE Transactions on Communications*, *COM-26(1)*, 215–227.

Simon, M. K., and W. C. Lindsey (1977), Optimum Performance of Suppressed Carrier Receivers, *IEEE Transactions on Communications*, *COM-25(2)*, 215–227.

Smith, A. M., C. N. Mitchell, R. J. Watson, R. W. Meggs, P. M. Kintner, K. Kauristie, and F. Honary (2007), GPS scintillation in the high arctic associated with an auroral arc, *Space Weather*, in press.

Sobel, D. (1995), *Longitude: the true story of a lone genius who solved the greatest scientific problem of his time*, Penguin Books, New York.

Sojka, J. J., D. Rice, V. Eccles, T. Berkey, P. Kintner, and W. Denig (2004), Understanding midlatitude space weather: Storm impacts observed at Bear Lake Observatory on 31 March 2001, *Space Weather*, *2*, doi:10.1029/2004SW000086.

Spilker, J. J. (1996a), *Global Positioning System: Theory and Applications*, vol. 1, chap. 3: Signal Structure and Theoretical Performance, pp. 57–119, American Institute of Aeronautics and Astronautics, Washington, D.C.

Spilker, J. J. (1996b), *Global Positioning System: Theory and Applications*, vol. 1, chap. 13, Tropospheric Effects on GPS, pp. 517–546, American Institute of Aeronautics and Astronautics, Washington, D.C.

Spilker, J. J. (1996c), *Global Positioning System: Theory and Applications*, vol. 1,

chap. 4, GPS Navigation Data, pp. 121–176, American Institute of Aeronautics and Astronautics, Washington, D.C.

Spilker, J. J., and F. D. Natali (1996), *Global Positioning System: Theory and Applications*, vol. 1, chap. 20, Interference Effects and Mitigation Techniques, pp. 717–771, American Institute of Aeronautics and Astronautics, Washington, D.C.

Spilker, J. J., and B. W. Parkinson (1996), *Global Positioning System: Theory and Applications*, vol. 1, chap. 2: Overview of GPS Operation and Design, pp. 29–55, American Institute of Aeronautics and Astronautics, Washington, D.C.

Stansell, T. A. (1983), The TRANSIT Navigation Satellite System, *Magnavox Technical Report, R-5933A*.

Strickland, D. J., R. E. Daniel Jr., and J. D. Craven (2001), Negative ionospheric storm coincident with DE1-observed thermospheric disturbance on October 14, 1981, *J. Geophys. Res.*, *106*.

Thomson, N. R., C. J. Rodger, and M. A. Clilverd (2005), Large solar flares and their ionospheric *D* region enhancements, *J. Geophys. Res.*, *110*, doi:10.1029/2005JA011008.

Townsend, B. R., P. C. Fenton, K. J. Van Dierendonck, and D. J. R. van Nee (1995), L1 carrier phase multipath error reduction using MEDLL technology, in *Proceedings of the 8th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GPS 1995*, pp. 1539–1544, Institute of Navigation.

Treumann, R. A. (2006), The electron cyclotron maser for astrophysical ap-

plication, *Annual Reviews of Astronomy and Astrophysics*, *13*, 229–315, doi: 10.1007/s00159-006-0001-y.

Tsunoda, R. T. (1985), Control of the seasonal and longitudinal occurrence of equatorial scintillations by the longitudina gradient in integrated E region Pedersen conductivity, *J. Geophys. Res.*, *90*, 447–456.

Tsurutani, B. T., W. D. Gonzalez, G. S. Lakhina, and S. Alex (2003), The extreme magnetic storm of 1-2 September 1859, *J. Geophys. Res.*, *108*(A7), doi:10.1029/2002JA009504.

Tsurutani, B. T., D. L. Judge, F. L. Guarnieri, P. Gangopadhyay, A. Jones, J. Nuttall, G. A. Zambon, L. Didkovsky, A. J. Mannucci, B. Iijima, R. R. Meier, T. J. Immel, T. N. Woods, S. Prasad, L. Floyd, J. Huba, S. C. Solomon, P. Straus, and R. Viereck (2005), The October 28, 2003 extreme EUV solar flare and resultant extreme ionospheric effects: Comparison to other Halloween events and the Bastille Day event, *Geophys. Res. Lett.*, *32*, doi:10.1029/2004GL021475.

Valladares, C. E., J. Villalobos, M. A. Hei, R. Sheehan, S. Basu, E. MacKenzie, and P. H. Doherty (2007), Simultaneous observation of travelling ionospheric disturbances in the northern and southern latitudes, *Ann. Geophys.*, submitted.

Van Dierendonck, A. J. (1996), *Global Positioning System: Theory and Applications*, chap. 8: GPS Receivers, pp. 329–407, American Institute of Aeronautics and Astronautics, Washington, D.C.

Van Dierendonck, A. J. (1998), How GPS receivers measure (or should measure) ionospheric scintillation and how GPS receivers are affected by the ionosphere.

Ward, P. (1996a), *Understanding GPS Principles and Applications*, chap. 5: Satel-

lite Signal Acquisition and Tracking, pp. 119–208, Artech House Publishers, Boston.

Ward, P. (1996b), *Understanding GPS Principles and Applications*, chap. 6: Effects of RF Interference on GPS Satellite Receiver Tracking, pp. 209 – 236, Artech House Publishers, Boston.

Ward, P. W. (1998), Performance Comparisons Between FLL, PLL and a Novel FLL-Assisted-PLL Carrier Tracking Loop Under RF Interference Conditions, in *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GPS 1998*, pp. 783–795, The Institute of Navigation.

Weill, L. R. (2006), Theoretical and practical sensitivity limits for assisted GNSS receivers using legacy and future GNSS signals, in *Proceedings of the 19th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2006*, pp. 2930–2943, The Institute of Navigation.

Woo, K. T. (2000), Optimum semicodeless carrier-phase tracking of L2, *Navigation, 47(2)*, 82–99.

Woon, J., T. Pany, and G. W. Hein (2006), GNSS software defined radio: real receiver or just a tool for experts?, *Inside GNSS*, pp. 48–56.

Yeh, K. C., and C. H. Liu (1982), Radio wave scintillations in the ionosphere, *Proc. IEEE, 70*, 324–360.